Machine Learning- and Model-based Image Fusion for cardiac interventions

Toth, Daniel

Awarding institution:
King's College London

KING'S COLLEGE LONDON

Machine Learning- and Model-based Image Fusion for Cardiac Interventions

Author:

Daniel TOTH

Supervisors:

Prof. Dr. Kawal RHODE

Prof. Dr. Aldo RINALDI

Dr. Peter MOUNTNEY

A thesis submitted in fulfillment of the requirements

for the degree of ***Doctor of Philosophy***

at the

School of Biomedical Engineering and Imaging Sciences

King's College London

2019

"In science, it is not speed that is the most important. It is the dedication, the commitment, the interest and the will to know something and to understand it – these are the things that come first."

Jenő Wigner

Abstract

In cardiac interventions, such as cardiac resynchronization therapy (CRT), fluoroscopy guidance can be enhanced through overlaying information extracted from preoperative magnetic resonance imaging (MRI) data. Registering multimodal image data, such as 3D/2D cine MRI to X-ray, however, remains a major research challenge. Due to fundamental differences in the image acquisition physics of MRI and X-ray fluoroscopy, no similar intensities or shared features are available. Due to the lack of shared features between the modalities, the use of classical intensity- or feature-based registration approaches is not feasible. This thesis proposes two main approaches to address this problem: 1) adjacent anatomical model-based registration and 2) imitation learning-based model-to-image registration. The adjacent anatomy-based approach relies on extracted models of the left ventricle (LV) from MRI and a reconstructed point cloud of the coronary veins from two interventional, contrasted X-rays (venograms). The method exploits the anatomical adjacency of the LV and the coronary veins through a globally optimal point cloud registration: globally optimal iterative closest point (GO-ICP). The approach has demonstrated high robustness and accuracy on phantom and clinical patient data. However, the approach is greatly dependent on the quality of acquired venograms. To create a more generic approach, an imitation learning-based method is proposed that is able to register the LV model to a single, non-contrast-enhanced X-ray acquisition. An artificial neural network (ANN) predicts transformations of the 3D LV model iteratively, relative to the 2D X-ray image to register to. The training of the ANN is performed entirely on digitally reconstructed radiographs (DRRs), artificial X-ray data generated from computed tomography (CT) volumes. The approach provides high accuracy on DRRs and high robustness on clinical X-ray data, however, results on the target domain are highly training parameter-dependent. The influence of parameters, i.e., training data order and network weight initialization, is investigated extensively. It is shown that domain randomization, i.e., applying unrealistic

perturbations to the synthetic training data, can result in substantially more consistent robustness on the target domain. The imitation learning-based registration approach was integrated into a clinical prototype for the interventional guidance of CRT procedures. After initial, successfully guided cases, the approach is to be extensively validated in a multicenter clinical trial.

Acknowledgements

First of all, I would like to thank my advisors, Prof. Dr. Kawal Rhode, Dr. Peter Mountney, and Prof. Dr. Aldo Rinaldi.

I would like to thank Kawal for all the support at King's and the freedom he has given me in pursuing my research. Thank you for being always available, creating such a casual environment, and for supporting me in every possible way at King's. Kawal has been providing guidance, feedback on my papers and reports, and helping me to find the right ways in the bureaucratic system.

Pete has been an excellent industrial advisor and line manager throughout the work towards this thesis. I cannot thank him enough for believing in me and all the support that I have received from him on all fronts. Pete has pushed me to achieve greater things than I ever thought were possible. Thank you for the countless hours discussing ideas and providing feedback on my papers and this thesis.

I would also like to thank Aldo for the support as a clinical advisor and as the clinical lead of the CRT project. The discussions about the clinical requirements have shaped my work significantly. Furthermore, the acquisition of the data used in this thesis could not have been possible without Aldo and the clinical registrars who have supported our procedures: Jonathan Behar, Ben Sieniewicz, Justin Gould and Bal Sidhu. Big thanks go to all of them.

I would also like to thank my thesis progression committee, Prof. Dr. Julia Schnabel, Dr. Andy King and Prof. Dr. Mark O'Neill for providing me with feedback and having such a casual atmosphere in the meetings.

I would also like to thank Dr. Tanja Kurzendorfer, Dr. Alexander Brost, and Dr. Martin Ostermeier from Siemens AT in Forchheim and Dr. Craig Buckley from Siemens UK. Alex has offered me to pursue my PhD at King's and I will never forget his first question: "Would you mind not getting a Dr.-Ing.?". Thank you to Tanja and Alex for supporting me scientifically, having fruitful discussions and reviewing my papers. Furthermore, as managers of the CRT guidance project and Martin as the group lead. I also have to thank Craig for being my line

manager initially and being so responsive to the bureaucratic challenges I have faced. I am very grateful for the support and the financial funding that I have received from Siemens through them, thus creating the financial foundations of this work.

I also thank Dr. Tommaso Mansi from Siemens MIT, Princeton for being so open to collaboration and for Shun Miao's advice and supervision. Thank you for all the support I have received, the collaboration, and the fun that I have had with the guys from Princeton: Sebastien Piat, Pascal Ceccaldi, Paul Klein, Serkan Cimen, Abdoul Amadou, and Walid Bekthaoui. The second half of my PhD would have been a different experience without them.

Special thanks go to everyone who has reviewed or proofread this work: Tanja, Nikolaus, Pascal, Dóri, and Dóra. Their work is much appreciated. Their feedback has made this thesis more professional and a more enjoyable read.

Thanks also go to everyone who has been sitting with me in the 4th Floor, North Wing office throughout the years: Sabrina, Sebastien, Pascal, Paul, Abdoul, Walid, Serkan, Luca and Nikolaus from Siemens and Maria, Sri, Mazen, James, Rashed, Shuangyi, Shu, and Yimin from King's. It has been great fun and thanks for the countless lunches we've had together.

The most special thanks go to my loved ones: my mom Mária, my dad János, my sister Dóra, and, in particular, my girlfriend Dóra. I could have never gotten so far without them believing in me and the continuous support that I have received from them, regardless of what my aims were. Dóri, thank you for being patient with me when I have worked more than I probably should have and that I could not spend as much time with you as you would have deserved. I hope, I can spend the rest of my life with you!

Contents

List of Figures

List of Tables

List of Acronyms

CO_2	carbon dioxide.
4CH	four chamber.
AFib	atrial fibrillation.
AHA	American Heart Association.
AI	artificial intelligence.
ANN	artificial neural network.
AP	anterior-posterior.
AV	aortic valve.
BEP	bull's eye plot.
BnB	branch and bound.
CAD	computer assisted design.
CAU	caudal.
CBCT	cone beam computed tomography.
CC	cross correlation.
CDF	cumulative distribution function.
CL	centerline.
CNN	convolutional neural network.
CPD	coherent point drift.
CR	clinical range.
CRA	cranial.
CRT	cardiac resynchronization therapy.
CS	coronary sinus.
CT	computed tomography.
CTA	computed tomography angiography.

Cycle-GAN cycle consistent GAN.

DOF degrees of freedom.

DR domain randomization.

DRN domain randomized network.

DRR digitally reconstructed radiograph.

EAM electro-anatomical mapping.

ECC entropy correlation coefficient.

ECG electrocardiogram.

EP electrophysiology.

FOV field of view.

FPS frames per second.

GAN generative adversarial network.

GC gradient correlation.

GI gradient information.

GO gradient orientation.

GO-ICP globally optimal iterative closest point.

GR global range.

GT ground truth.

HF heart failure.

HU Hounsfield unit.

ICP iterative closest point.

IRED infrared light emitting diode.

KLD Kullback-Leibler divergence.

LA left atrium.

LAO left anterior oblique.

LAx long axis.

LBBB left bundle branch block.

LGE late gadolinium enhancement.

LV left ventricle.

MAE mean absolute error.

MI mutual information.

ML machine learning.

MR magnetic resonance.

MRI magnetic resonance imaging.

MV mitral valve.

NCC normalized cross correlation.

NN neural network.

NYHA New York Heart Association.

PET positron emission tomography.

RA right atrium.

RAO right anterior oblique.

ReLU rectified linear unit.

RMS root mean square.

ROI region of interest.

RV right ventricle.

RXA rotational X-ray angiography.

SAx short axis.

SIFT scale invariant feature transform.

SN standard network.

SPECT single photon emission computed tomography.

SSD sum of squared differences.

SSFP steady state free precession.

SVD singular value decomposition.

TEE transesophageal echocardiography.

TRE target registration error.

TTE transthoracic echocardiography.

US ultrasound.

XA X-ray angiography.

XMR MR/X-ray suite.

Chapter 1

Introduction

Patients with advanced drug-refractory heart failure (HF) can be safely treated with cardiac resynchronization therapy (CRT). However, 30 to 50 % of patients do not respond to the treatment [1]. In this procedure, a CRT device is implanted using fluoroscopic image guidance. The device has three leads which are placed in the right atrium (RA), right ventricle (RV), and through the coronary sinus (CS) on the surface of the left ventricle (LV). Suboptimal placement of the lead on the LV has been identified as a leading cause of non-response. Unfortunately, improving the placement of this lead is extremely challenging for clinicians.

It has been shown that placing the LV lead away from scar tissue and at the site of the latest point of mechanical activation can improve the response rate [2]. This information cannot be directly obtained from fluoroscopic images and requires an additional imaging modality, such as preoperative magnetic resonance imaging (MRI). Registering preoperative MRI with intraoperative fluoroscopic images enables clinicians to visualize scar tissue and mechanical activation in real time fluoroscopic overlay, guiding the placement of the LV lead to improve response rates.

Registering MRI to fluoroscopy remains an open research problem. The main challenge is the lack of shared information between the modalities. In the case of cardiac MRI, the images show high soft tissue contrast to visualize the heart, but bony structures, such as the spine, are not easily seen. In fluoroscopic images however, bony structures and instruments are easily visible, but there is a lack of contrast for soft tissue. The only way to visualize soft tissue anatomy is to

inject contrast agent. In CRT this is used to visualize the CS [3]. The registration problem is further complicated by different fields of view, the low resolution of MRI, and cardiac and respiratory motions.

Although clinical experts are capable of registering a mesh model to the X-ray images, the process is time-consuming and the results are operator-dependent. The manual registration process can be supported through the application of fiducial markers to the skin surface of the patient before the preoperative scan [4]. However, the markers have to remain on the patient for the intervention and this is often not realistic, due to the preoperative scan mostly being acquired not directly before the procedure, often weeks in advance. Furthermore, a fiducial-based registration, does not guarantee a sufficiently accurate registration of the heart [5]. Therefore, an automatic registration approach is desired.

Previous approaches registering pre- to intraoperative data for cardiac procedures either disturbed the clinical workflow or relied on too many, somewhat unrealistic assumptions. Methods have been developed to use tools [6, 7, 8], however, tools may deform the anatomy. Approaches have emerged that utilise more advanced MRI protocols, however, these are often not feasible for HF patients [9]. Other approaches have made assumptions and involved a significant amount of manual interaction to perform the registration [10, 11].

This work proposes registration methods to address the problem of the lack of cross-modality landmarks between cardiac cine magnetic resonance (MR) images and contrast enhanced X-ray fluoroscopy. The overall aim was to not require fiducial markers, novel MRI sequences, or rely on tools, thereby enable the developed approaches to seamlessly integrate into the current clinical workflow. Two main approaches, and their variations, are presented to provide solutions to this challenging registration problem.

1.1 Contributions

The objective of this work is to explore novel registration approaches for image guided cardiac procedures, especially for guidance of CRT procedures with

MRI data, having minimal interference with the clinical workflow. The main contributions of this thesis are:

- An adjacent anatomical model-based framework was developed for the registration of the 3D reconstruction of the coronary veins from two X-ray angulations to a LV mesh model, extracted from MR. The registration is performed in 3D by a classical optimization-based approach, globally optimal iterative closest point (GO-ICP), that is a combination of branch and bound (BnB) and iterative closest point (ICP).

- The framework was extended by performing a semiautomatic segmentation of the coronary veins in the contrasted X-ray images and reconstructing every possible epipolar correspondence (superabundant reconstruction). The adjacent anatomical structures are registered with a version of the GO-ICP method with outlier rejection.

- A method was developed for the evaluation of cardiac registration by reconstructing artificial valves from the X-ray projections and comparing their positions with theirs extracted from the MR images.

- An imitation learning-based registration framework was developed for registering cardiac models to interventional X-ray images. The method is trained on fully synthetic data and does not rely on contrast agent injections. It is capable of registering the LV model to a single, non-contrasted X-ray image.

- The imitation learning framework was extended to multiple angulations, i.e., right anterior oblique (RAO) and left anterior oblique (LAO) to realize feasibility for electrophysiology (EP) procedures, i.e., CRT.

- It was shown that training with different parameters, i.e., network weight initialization or training data order, can result in low and inconsistent performance on the target domain of clinically acquired X-ray data.

- Domain randomization (DR) was performed to realize a more efficient and consistent transfer to increase the robustness on the clinical CRT data in the imitation learning-based framework.

1.2 Thesis Outline

In Chapter 2, the medical background of this work is highlighted. The anatomy of the human heart, with particular focus on the LV and the CS, and CRT is described. Details of CRT are given, with focus on the delivery of the therapy in a catheterization laboratory and the inherent challenges thereof.

Chapter 3 provides a thorough literature review. A critical analysis of state-of-the-art registration approaches, is presented. The review has particular focus on registration methods relevant for cardiac procedures, especially for CRT. To further motivate the work, an experiment to measure manual registration performace is presented.

In Chapter 4, a method exploiting prior knowledge of adjacent anatomical structures is presented. Preoperative MR images are segmented to extract a 3D model of the LV and a 3D model of the CS is reconstructed from two contrasted intraoperative fluoroscopic images. The two models of adjacent anatomy are registered using a globally optimal technique well suited to registering partial surfaces. The method is extended to use a reconstruction of the vessels without known correspondences, named superabundant reconstrucion. A version of the globally optimal technique with outlier removal is used to perform the registration.

Chapter 5 describes an imitation learning-based approach, registering a segmented LV from MR, to uncontrasted X-ray images. Manual feature engineering is not required. The method relies on a neural network-based framework, able to automatically extract features and make predictions based on them. This results in a system that is able to utilize the highly limited amount of cross-modality information between the two modalities. Due to the lack of MR and X-ray data pairs and of ground truth (GT) registrations, training was performed on fully

synthetic data. Only computed tomography (CT) data are used to generate the synthetic training data. LV mesh models can be extracted from the CTs and they can be used to generate digitally reconstructed radiograph (DRR) projection images to create artifical X-rays. The generated synthetic data is perfectly registered, thus a GT is available.

In Chapter 6, the transfer performance to the target domain of clinically acquired X-ray data is evaluated for varying training parameters, i.e., network weight initialization and training data order. It is shown that performance for different training parameters can suffer greatly and be inconsistent, without explicit transfer learning measures implemented. DR is performed, i.e., random perturbations are applied to the training data to generate unrealistic images. This increases the probability of a successful transfer, i.e., increased and more consistent performance on the target domain, because the data of the target domain is perceived by the imitation learning agent as just another variation of the possible images.

The thesis is concluded with Chapter 7, giving a brief conclusion and highlighting the potential future course of research in the field of image registration for cardiac procedures, with special account for the clinical applications, in particular CRT.

1.3 Resulting Publications

The following publications have resulted from the work presented in this thesis:

> D. Toth, M. Panayiotou, A. Brost, J. M. Behar, C. A. Rinaldi, K. S. Rhode, and P. Mountney. “Registration with adjacent anatomical structures for cardiac resynchronization therapy guidance”. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Ed. by T. Mansi, K. McLeod, M. Pop, K. Rhode, M. Sermesant, and A. Young. Vol. 10124 LNCS. Athens, Greece: Springer, Cham, 2017, pp. 127–134. DOI: 10.1007/978-3-319-52718-5_14.

D. Toth, M. Panayiotou, A. Brost, J. M. Behar, C. A. Rinaldi, K. S. Rhode, and P. Mountney. "3D/2D Registration with superabundant vessel reconstruction for cardiac resynchronization therapy". In: *Medical Image Analysis* 42 (2017), pp. 160–172. DOI: 10.1016/j.media.2017.08.001.

D. Toth, S. Miao, T. Kurzendorfer, C. A. Rinaldi, R. Liao, T. Mansi, K. Rhode, and P. Mountney. "3D/2D model-to-image registration by imitation learning for cardiac procedures". In: *International Journal of Computer Assisted Radiology and Surgery* 13.8 (2018), pp. 1141–1149. DOI: 10.1007/s11548-018-1774-y.

D. Toth, S. Cimen, P. Ceccaldi, T. Kurzendorfer, K. Rhode, and P. Mountney. "Training Deep Networks on Domain Randomized Synthetic X-ray Data for Cardiac Interventions". In: *Medical Imaging with Deep Learning*. Submitted, 2019, pp. 1–10.

M. Panayiotou, P. Mountney, A. Brost, D. Toth, T. Jackson, J. M. Behar, C. A. Rinaldi, R. J. Housden, and K. S. Rhode. "Dynamic mapping of ventricular function from cardiovascular magnetic resonance imaging". In: *38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. Orlando, Florida, USA, 2016, pp. 4137–4140. DOI: 10.1109/EMBC.2016.7591637.

S. Reiml, D. Toth, M. Panayiotou, B. Fahn, R. Karim, J. M. Behar, C. A. Rinaldi, R. Razavi, K. S. Rhode, A. Brost, and P. Mountney. "Interactive visualization for scar transmurality in cardiac resynchronization therapy". In: *Proc. SPIE 9786, Medical Imaging 2016: Image-Guided Procedures, Robotic Interventions, and Modeling. Vol. 9786*. San Diego, California, USA: SPIE, 2016, 97862S. DOI: 10.1117/12.2214737.

J. M. Behar, P. Mountney, D. Toth, S. Reiml, M. Panayiotou, A. Brost, B. Fahn, R. Karim, S. Claridge, T. Jackson, B. Sieniewicz, N. Patel, M. O'Neill,

R. Razavi, K. S. Rhode, and C. A. Rinaldi. "Real Time X-MRI Guided Left Ventricular Lead Implantation for Targeted Delivery of Cardiac Resynchronization Therapy". In: *JACC: Clinical Electrophysiology* 3.8 (2017), pp. 803–814. DOI: 10.1016/j.jacep.2017.01.018.

J. M. Behar, B. Sieniewicz, P. Mountney, D. Toth, M. Panayiotou, S. Claridge, K. Rhode, and C. A. Rinaldi. "Image Integration to Guide Wireless Endocardial LV Electrode Implantation for CRT". In: *JACC: Cardiovascular Imaging* 10.12 (2017), pp. 1526–1528. DOI: 10.1016/j.jcmg.2017.01.015.

P. Mountney, J. M. Behar, D. Toth, M. Panayiotou, S. Reiml, M. P. Jolly, R. Karim, L. Zhang, A. Brost, C. A. Rinaldi, and K. Rhode. "A Planning and Guidance Platform for Cardiac Resynchronization Therapy". In: *IEEE Transactions on Medical Imaging* 36.11 (2017), pp. 2366–2375. DOI: 10.1109/TMI.2017.2720158.

M. Panayiotou, D. Toth, T. Adem, P. Mountney, A. Brost, J. M. Behar, C. A. Rinaldi, R. J. Housden, and K. S. Rhode. "3D reconstruction of coronary veins from a single X-ray fluoroscopic image and pre-operative MR". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 10124 LNCS. Athens, Greece: Springer, 2017, pp. 66–75. DOI: 10.1007/978-3-319-52718-5_8.

S. Reiml, T. Kurzendorfer, D. Toth, P. Mountney, M. Panayiotou, J. M. Behar, C. A. Rinaldi, K. Rhode, A. Maier, and A. Brost. "Automatic layer generation for scar transmurality visualization". In: *Bildverarbeitung für die Medizin 2017. Informatik aktuell*. Ed. by K. H. Maier-Hein, G. Fritzsche, T. M. Deserno, G. Lehmann, H. Handels, and T. Tolxdorff. Springer Vieweg,

Berlin, Heidelberg, 2017, pp. 167–172. DOI: 10.1007/978-3-662-54345-0_40.

S. Reiml, T. Kurzendorfer, D. Toth, P. Mountney, S. Steidl, A. Brost, and A. Maier. "Automatic Vertebrae Segmentation in Fluoroscopic Images for Electrophysiology". In: *2017 IEEE Nuclear Science Symposium and Medical Imaging Conference*. 2017, pp. 1–3.

M. Alhrishy, D. Toth, S. A. Narayan, Y. Ma, T. Kurzendorfer, K. Rhode, and P. Mountney. "A Machine Learning Framework for Context Specific Collimation and Workflowphase Detection". In: *15th International Symposium on Computer Methods in Biomechanics and Biomedical Engineering*. Ed. by P. R. Fernandes and J. M. Tavares. Lisboa, Portugal, 2018, pp. 1–10.

R. Karim, M. Panayiotou, O. Chowdhury, R. J. Housden, S. Hummady, D. Toth, T. Kurzendorfer, P. Mountney, and K. Rhode. "Image Data Analysis for Quantifying Scar Transmurality in MRI phantoms for Cardiac Resynchronisation Therapy". In: *40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. Honolulu, HI, USA, 2018, pp. 1111–1114. DOI: 10.1109/EMBC.2018.8512448.

E. Thomas, D. Toth, T. Kurzendorfer, K. Rhode, and P. Mountney. "Mechanical Activation Computation from Fluoroscopy for Guided Cardiac Resynchronization Therapy". In: *40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. Honolulu, HI, USA, 2018, pp. 592–595. DOI: 10.1109/EMBC.2018.8512434.

Chapter 2

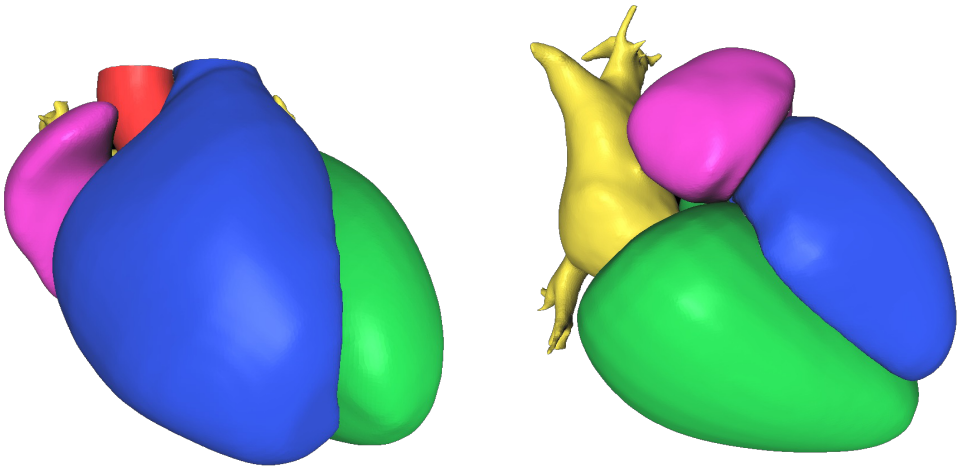
Medical Background

In this chapter, the medical background of this work is described in detail. First, the anatomy and physiology of the heart are presented, with particular focus on the left ventricle (LV) and its main vein, the coronary sinus (CS). Second, the target clinical application, cardiac resynchronization therapy (CRT), and the challenges involved therein are highlighted.

2.1 The Human Heart

The human heart is the central organ for the circulation of blood, the oxygenation of the organs in the human body. It consists of four chambers, see Figure 2.1, that have their own characteristics and responsibilities. The LV supplies the body with oxygen-rich blood through the aortic valve (AV), the aorta, and the arteries. The deoxygenated, carbon dioxide (CO_2) rich blood returns through the venous system into the right atrium (RA), to be forwarded through the tricuspid valve into the right ventricle (RV). The RV pumps it to the lungs, to exchange the CO_2 for oxygen. The oxygen-rich blood returns to the left atrium (LA) and is let through the mitral valve (MV) to the LV to complete the circle [12].

In the following sections, the most relevant anatomical structures for CRT, i.e., the LV and the CS, are described in more detail.



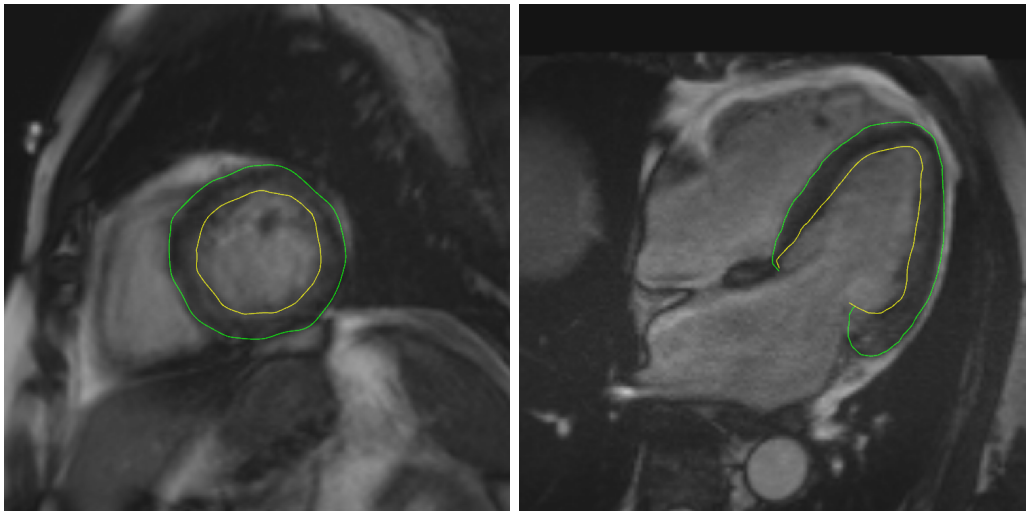
(a) AP view (b) Inferior view

Figure 2.1: Schematic illustrations of the chambers of the heart: LV (green), RV (blue), LA (yellow), RA (purple), and the root of the aorta (red). Shown in (a) anterior-posterior (AP) and (b) inferior radiographic orientations.

2.1.1 Left Ventricle

The LV is responsible for supplying the body with oxygenated blood through the aorta and arteries. It mainly consists of muscle tissue, myocardium, that contracts and relaxes periodically. The contraction of the myocardium is controlled by the electrical system of the heart. The electrical impulses originate from the sino-atrial node, which is situated in the upper part of the RA and serves as the natural pacemaker of the heart. The electrical signal travels through the atria, inducing contraction, the atrioventricular node, and the bundle of His. The signal separates into two branches, the right and the left bundles. The two bundles excite the myocardium through thin embedded fibers, the Purkinje fibers [13]. The contraction of the muscles results in the ejection of blood from the LV cavity.

The LV wall has three layers. The main layer, as mentioned before, is the myocardium, the cardiac muscle tissue. The myocardium is separated from the blood pool in the inside of the ventricle by the endocardium and it is bound by the epicardium on the outer surface, see Figure 2.2.



(a) SAx view (b) 4CH view

Figure 2.2: Magnetic resonance (MR) acquisitions of the LV with the endo- (yellow) and epicardium (green) highlighted. (a) Short axis (SAx) acquisition slice of the LV. (b) Four chamber (4CH) acquisition slice of the LV.

2.1.2 Coronary Sinus

The CS is the main coronary vein of the left ventricle. It collects deoxygenated blood from the myocardium through its tributaries and drains into the RA. Its tributaries are the great cardiac vein, the left marginal vein, the posterior cardiac vein, and the middle cardiac vein. Unlike the coronary arteries', the coronary venous anatomy has high variation throughout patients. Some of the mentioned tributaries are often not present and their location may vary too [14]. Exemplary contrasted, interventional X-ray angiography (XA) acquisitions, so called venograms, of the CS are shown in Figure 2.3.

2.2 Cardiac Resynchronization Therapy

Patients with moderate to severe symptoms of heart failure (HF) can be treated with CRT. In CRT a biventricular pacemaker is implanted to resynchronize the ventricles of the patient, thus improve ventricular function. Successful CRT reduces the symptoms of heart failure, which provides patients with improved quality of life and results in lower mortality [15, 16, 17].

(a) LAO (b) AP (c) RAO

Figure 2.3: Contrasted XA acquisitions of the CS during an intervention. Projections of (a) left anterior oblique (LAO) 30°, (b) anterior-posterior (AP), and (c) right anterior oblique (RAO) 30° angulations.

In the next sections, HF, CRT delivery, its challenges, and a clinical prototype to guide CRT implantations are described in detail.

2.2.1 Heart Failure

The term HF is used to describe the condition when the heart cannot supply the body with sufficient quantities of blood, thus oxygen. The most frequent causes of HF are hypertension and coronary disease [18]. Coronary disease can lead to a myocardial infarction. In an infarction, a coronary artery supplying the myocardium is blocked, resulting in insufficient blood supply of the cardiac tissue. In consequence, the affected myocardial tissue is not oxygenated adequately, resulting in myocardial tissue death. The dead myocardium undergoes scarring, resulting in a loss of electrical conductivity. This results in an impairment of electrical activation of the myocardium, often a left bundle branch block (LBBB), where the activation of the left ventricle is delayed. This can result in inter- (delay between ventricles) or intraventricular (delay within LV) dyssynchrony and could be the cause of other arrhythmias [19].

The diagnosis of heart failure is performed in the first line by evaluating the electrocardiogram (ECG) of the patient [20]. However, it has been shown that there is no direct correlation between HF and an elongated QRS complex [21, 22]. The diagnosis is confirmed by transthoracic ultrasound (US) imaging. US can be

used to identify the cardiac wall area of delayed activation, chamber volumes (end systolic/diastolic volume, ejection fraction) can be quantified, and blood flow can be measured through doppler imaging [23].

Severity of heart failure is classified by the New York Heart Association (NYHA) criteria into four classes, I to IV [24]. For patients of classes III and IV, moderate to severe HF, a biventricular pacemaker, a CRT device is advised [25]. It has been shown that CRT can achieve positive outcomes, i.e., increase in LV function, even in NYHA class II patients [26].

2.2.2 CRT Delivery

Patients having heart failure with moderate to severe symptoms can benefit from CRT. In CRT, a biventricular pacemaker is implanted to resynchronize an asynchronously beating left ventricle. While a conventional pacemaker has two leads, one in the RA and one in the RV, a biventricular pacemaker, or so called CRT device, has an additional third lead on the LV.

The implantation of the CRT device is a minimally invasive procedure that is performed under fluoroscopy guidance in the catheterization laboratory, see Figure 2.4. The device is implanted under the skin, below the clavicle (collarbone). During the implantation, venous acces is gained through the subclavian vein to the superior vena cava that drains into the RA. One lead is positioned in the RA, the second through the tricuspid valve into the RV, and the third lead onto the surface of the LV through the CS, the main vein of the LV, draining into the RA, see Figure 2.5 [27].

The main radiographic projection orientations for implanting the leads, i.e., working angulations for CRT are AP, left anterior oblique (LAO) 30°, and right anterior oblique (RAO) 30°. Images of caudal (CAU) and cranial (CRA) angulations are rarely acquired. Exemplary venogram acquisitions are shown in Figure 2.3.

In an ideal case, the device is able to resynchronize an asynchronously beating heart and the patients heart undergoes reverse remodelling over time,

Figure 2.4: CRT delivery in a catheterization laboratory.

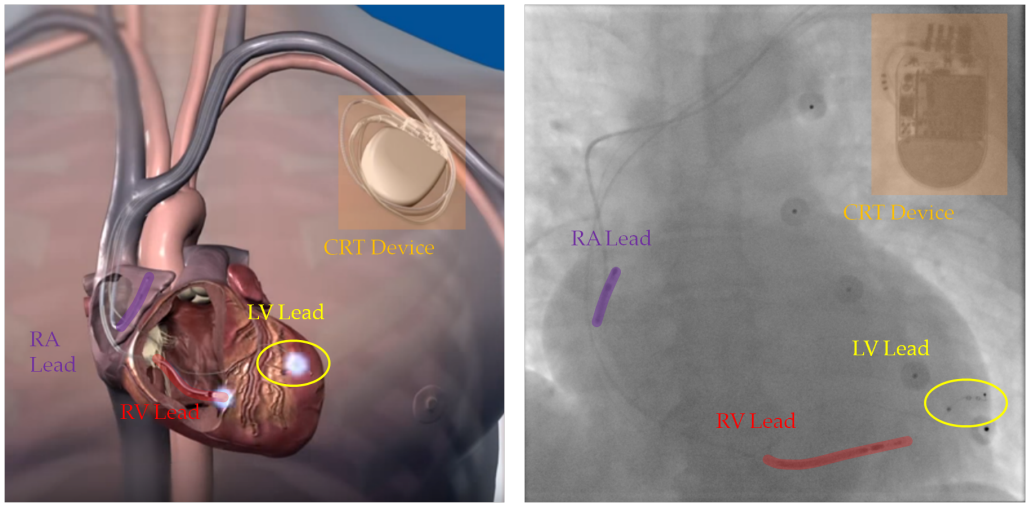
improving cardiac function, thus quality of life and life expectancy [29].

2.2.3 CRT Challenges

In current clinical practice, CRT is delivered with limited information, targeting the postero-lateral wall, since no soft tissue contrast is present in X-ray images. However, the optimal placement of the LV lead is patient specific [30]. Despite of undergoing CRT, 30 to 50 % of patients do not respond [2]. The high fail rate is associated with suboptimal placement of the LV lead.

The main factors in the success of pacing of the LV based on Mountney et al. [4] are as follows:

- Myocardial scar: Pacing in myocardial scar tissue is associated with poor outcomes, due to its low electrical conductivity [31]. Pacing away from areas of scar tissue results in better outcomes [32].



(a) Schematic illustration of CRT (b) X-ray image of CRT implant

Figure 2.5: Implanted biventricular pacemaker. (a) Schematic illustration of implanted CRT device (orange) and three electrodes in the RA (purple), RV (red), and on the LV (yellow) [28]. Courtesy of Medtronic. (b) X-ray view of implanted CRT device and electrodes.

- Mechanical activation: Pacing in areas of the myocardium showing the highest delays in mechanical activation can result in better reverse remodelling of the LV, thus improved clinical outcomes [33, 34].

- Coronary venous anatomy: The LV lead is lodged into one of the tributaries of the CS. The number and location of tributaries limits the potential targets. The target vein has to be sufficiently wide for the lead, but cannot be too wide for a secure placement.

- Anatomical position: Pacing on the antero- or postero-lateral wall was shown to lead to best outcomes [35, 36]. It is advised to place the lead away from the apex, to avoid proximity to the RV lead [37].

- Phrenic nerve stimulation: A further reason for avoiding an apical placement can be the unintended stimulation of the phrenic nerve [38]. If phrenic nerve stimulation is experienced during the intervention, the lead is retracted and placed into a more basal region of the LV or if a multipolar lead is implanted, it is electronically repositioned (reprogrammed) [39, 40].

It was shown in the studies presented above that, by controlling these factors, clinical outcomes of CRT can be improved. In current clinical practice involving X-ray guidance, only a highly limited amount of the necessary information is available. Venograms can be acquired of the coronary venous anatomy and phrenic nerve stimulation can be detected during the intervention. However, it might not occure in the supine position of the patient during the intervention. The other factors, i.e., mechanical activation and myocardial scar tissue position cannot be determined, although promising, early approaches exist [41].

In order to compensate for the lack of information in the LV lead placement, thus to improve success rates, preoperative information can be involved into the clinical decision making process [42, 43, 44]. Preoperative image data can be US [45, 46, 47], MR [48, 32, 49, 50], computed tomography (CT) [51], single photon emission computed tomography (SPECT) [52], or positron emission tomography (PET) [53]. However, none of the above can provide all the necessary information on their own. The possibly most promising approaches involve preoperative MR or CT imaging.

MR can provide the mechanical activation information and the extent and position of scar tissue [4]. It is, however, not possible to extract and visualize the coronary anatomy in standard cardiac protocols. More advanced acquisition protocols are often not feasible for HF patients.

CT can also provide the mechanical activation information, if multiphase images are acquired. The coronary venous anatomy can be imaged too. However, scar information is currently not derivable [51]. Since scar information is probably the most important factor in lead placement, MR is the more suitable modality for guiding CRT procedures at the current state of research.

Different ways exist to involve preoperative data into the interventional decision making process. Clinicians can visualize the preoperative acquisitions and make a decision prior to the procedure, mentally recalling where the target is intraoperatively. A mental projection of the 3D information into the 2D X-ray imaging plane has to be performed during the procedure. It is also possible

to have the pre- and intraoperative images side-by-side dispayed during the intervention, but the mental projection still needs to be performed. A more advanced approach is to fuse the pre- and intraoperative images and present the resulting images to the interventionalist during the procedure as an overlay [50, 4]. The overlay images are easier to work with, since the projection from 3D to 2D space is provided.

To facilitate such a solution, the pre- and intraoperative data have to be aligned, referred to as registration. Automatic registration of CT to X-ray [54] is feasible for various clinical applications, e.g., pedicle screw placement in the spine, relying on high contrast, bony structures [55]. However, in cardiac applications the target anatomy, i.e., the heart, consists of soft tissue. Registration by surrounding bony structures might not be accurate enough, especially due to deformations, relative motion of internal organs between the two acquisitions [5]. Registering MR to X-ray is even more challenging. The amount of shared information between the modalities is highly limited. The registration is performed mostly manually, however, it is challenging, due to the lack of cross-modality landmarks. A clinician of high experience can perform the registration, but the process is time-consuming and operator-dependent. Thus, to enable a valid overlay of preoperative information, an automatic registration method is required.

2.2.4 A Clinical Prototype for Guiding CRT

To target the challenges of CRT delivery, described in Section 2.2.3, a clinical prototype was designed for MR guidance (Siemens Healthcare GmbH, Forchheim, Germany) [4]. The prototype incorporates image processing components to support the whole clinical workflow, from the acquisition of the preoperative images, to the overlay of derived information onto the interventional X-ray images, see Figure 2.6.

First, the acquired cine MR images are segmented to extract mesh models of the endo- and epicardium. The acquired late gadolinium enhancement (LGE)

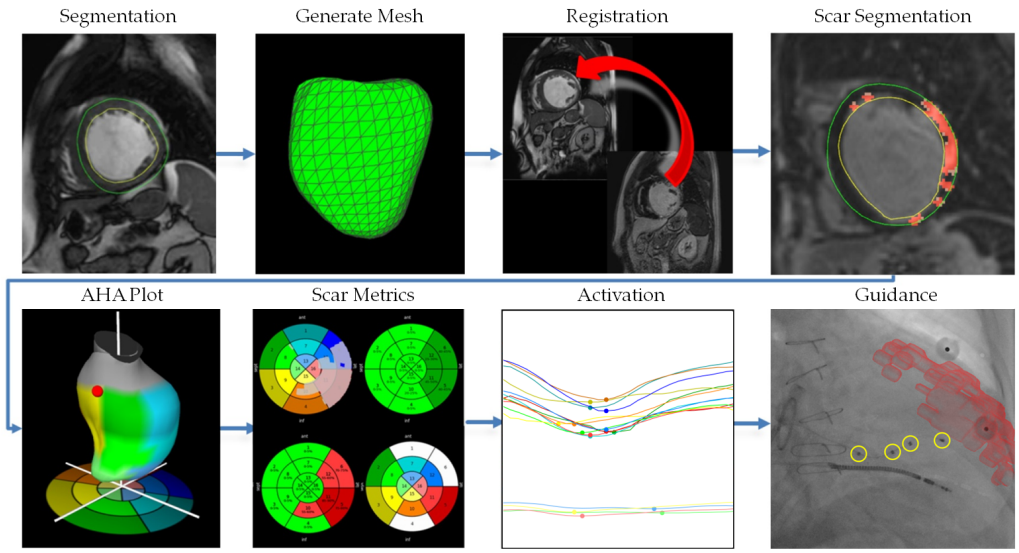
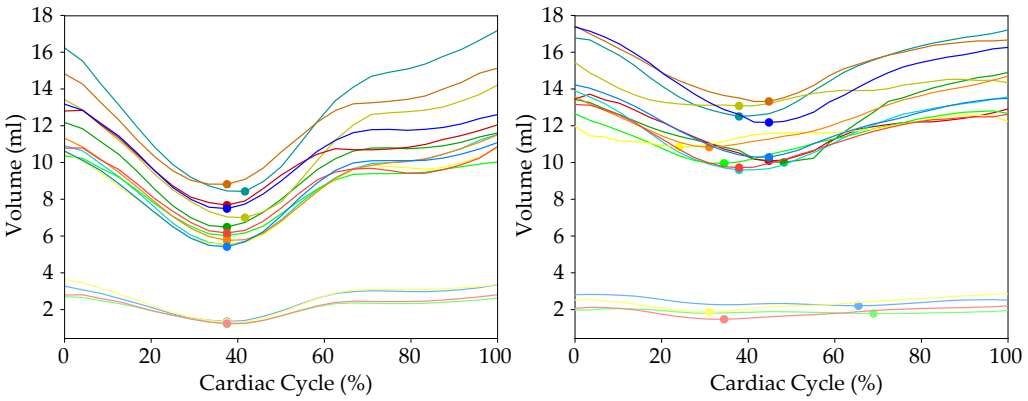


Figure 2.6: Image processing pipeline of the MR guided CRT prototype [4].

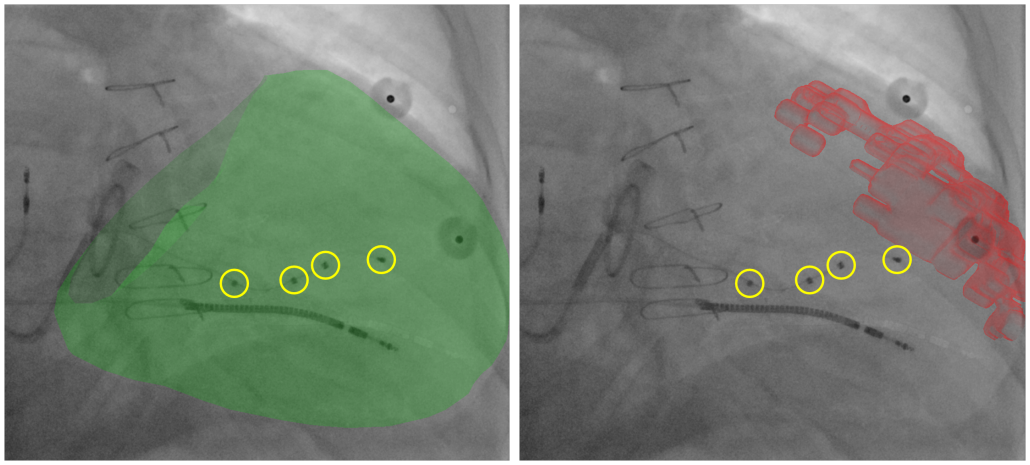
MR images for scar imaging are registered to the segmented cine images, to account for patient motion and deformations. Scar tissue is segmented and the corresponding American Heart Association (AHA) 16-segment subdivision is generated [56]. Based on the mapping, scar information is projected into a conventional 2D AHA plot. Mechanical activation is computed for each of the 16 segments individually from the cine magnetic resonance imaging (MRI) and visualized, see Figure 2.7. The presented information is used by the operator to identify potential targets. The targets or scar information can be overlaid after a successful manual registration between the mesh model of the LV and two angulations of the X-ray images. The overlay can be used to directly visualize soft tissue anatomy or scar tissue, in conjunction with the LV lead to be implanted, see Figure 2.8.

A valid overlay is challenging to achieve through 3D/2D manual registration. As described in Section 2.2.3, clinicians are able to register manually, but it is a user-dependent and time-consuming process. The main reason for this is the limited amount of cross-modality information. Having a robust and sufficiently accurate registration method available in a clinical prototype is much desired.



(a) Healthy volunteer (b) HF patient

Figure 2.7: Mechanical activation: Endocardial volume change of the LV over the cardiac cycle. Each line corresponds to a segment of the 16-segment AHA subdivision. The points annotate the global minima of the curves. (a) In the healthy volunteer, the minima are vertically aligned, the LV is in synchrony. (b) In a HF patient with LV dyssynchrony, the points are misaligned.



(a) LV model overlay (b) Scar tissue model overlay

Figure 2.8: Overlay images from the MR guided CRT clinical prototype (Siemens Healthcare GmbH, Forchheim, Germany). (a) Overlay of the segmented LV model (green) and the poles of a quadripolar lead highlighted (yellow). (b) Overlay of the segmented scar tissue (red) and the poles of the quadripolar lead highlighted (yellow).

2.3 The Future of Biventricular Pacing

There is a substantial amount of research in optimization of CRT procedures, to increase the success rates. One important method of making LV lead placement more successful is implanting leads with multiple electrode poles (bi- or quadripolar pacing), pacing positions along the lead [57], or placing leads into multiple veins simultaneously [58]. This has the benefit that not just a single position can be paced, but every pole along the lead or in each vein. A pattern of excitation, e.g., first and third pole pacing, can also be programmed during the implant and adjusted post-interventionally. It was shown that multipolar pacing is associated with lower mortality and elimination of phrenic nerve stimulation [59]. However, no substantial difference was shown compared to optimized (through scar analysis) monopolar lead placement [60].

Work was also done to realize preoperative data guided CRT delivery in cases where a standard pacemaker has been placed prior to biventricular pacing (non de novo implants). The main challenge is that implanted devices might not be MR compatible, thus MR guided pacing is not feasible. Even if previously implanted devices are MR compatible, the acquired MR image quality is mostly insufficient for targeted delivery. In these cases, a CT volume can be acquired and processed in a similar way as the MR images, to extract models and mechanical activation information. Scar analysis is, however, not available, due to limited contrast levels of scar tissue in CT images. The difference between myocardial tissue and scar tissue is in the order of the noise observed in Hounsfield unit (HU) values. Besides optimizing CT protocols to identify scar, current work focuses on the definition and targeting of the optimal pacing site, using the available information (mechanical activation and models) extended by electro-anatomical mapping (EAM) data [61, 62].

A further, radically different approach for CRT, endocardial pacing, is also showing promising results [63]. In endocardial pacing, a small, wireless electrode is implanted into the endocardium of the LV through arterial access to the chamber. Energy is transmitted to the electrode from an implanted battery,

through an US transducer. The benefit over conventional CRT is that the electrode can be placed freely on the inner surface of the LV. The pacing location is not restricted by the coronary venous tree anatomy. Superiority of endocardial over epicardial pacing was shown in terms of acute haemodynamic response [63]. The procedure, however, comes with larger trauma and infection risk through the surgical implantation of the battery, the US transducer, and the arterial, instead of venous access during the procedure [64].

Chapter 3

Registration for Image Guided Cardiac Interventions

In this chapter, the relevant literature for the registration of medical image data for the interventional guidance of cardiac resynchronization therapy (CRT) delivery is analyzed. In the first section, a brief overview of medical image registration in general is presented. Next, registration methods applicable for cardiac procedures, especially for CRT guidance, are summarized. Furthermore, an evaluation of manual registration performance between cardiac magnetic resonance (MR) and X-ray images is presented to motivate the development of automatic approaches. In the final section, a conclusion is given.

3.1 Medical Image Registration

The process of transforming multiple datasets, such as medical image data, into a common coordinate system is called registration. Medical image registration, or the registration of imaging systems, is one of the largest fields of medical image processing. It has been studied for decades and numerous methods were developed for various applications, such as disease progression monitoring [65] or image fusion for procedure guidance [66]. The breadth and depth of the field is marked by the number of medical image registration review articles and book chapters that were published throughout the years [67, 68, 69, 70, 71, 72, 73, 74].

There are several ways of categorizing medical image registration. Methods can be categorized into subfields based on the information used for registration (intensity vs. features), the number of modalities involved (inter- vs. intramodality), the desired transformation (rigid vs. nonrigid), or based on dimensionality of the datasets to be registered (e.g. 3D/2D or 3D/3D). In the following sections, these categories will be highlighted briefly. Since currently the focus of research is shifting towards machine learning-based approaches, a brief overview of learning-based methods is given.

3.1.1 Intensity- and Feature-based Registration

One way of registering medical images is to design and optimize a similarity measure based on pixel (or voxel) intensity values. A number of intensity-based metrics were developed for such purposes, such as sum of squared differences (SSD), cross correlation (CC), or normalized cross correlation (NCC) [75]. The benefit of intensity-based approaches is that no segmentation or feature detection is required, the algorithms can operate directly on the image intensity data.

This principle was extended by the introduction of measures based on entropy, such as mutual information (MI) [76] or the Kullback-Leibler divergence (KLD) [77]. Methods based on these metrics do not compare the pixelwise intensity values directly, but generate histograms of the images and optimize the entropy-based metric of the histograms (maximizing MI or minimizing KLD). An extensive review of MI-based approaches was published in [78].

In feature-based registration, corresponding points of interest are extracted from the images to register, to define a geometric transformation between them. The feature points and their correspondence can be defined manually by simple point selection or automatically by feature detection and matching, e.g., by scale invariant feature transform (SIFT) features [79]. However, automatic feature matching is more challenging in the case of some modalities, such as X-ray imaging. X-ray images are absorption images. If recorded at different angulations, the variation of the image data prevents accurate feature matching. Due to

common feature detectors being designed for reflection images, most common in computer vision.

3.1.2 Inter- and Intramodality Registration

If a registration algorithm operates on data from a single modality, such as computed tomography (CT) to CT registration, it is called an intramodality registration technique. However, if the registration is performed between multiple different modalities, such as CT to MR, it is referred to as intermodality registration.

In certain cases, intramodality registration can be solved by applying intensity based methods, e.g., in the case of CT to CT registration, because there is a consistency in intensity values: the same tissue type will have similar intensity, Hounsfield units (HUs), in the two images. If we consider MR to MR registration, the same structures might appear having different intensities, due to the various different acquisition protocols, such as T1 or T2 mapping [80].

In intermodality registration, intensity-based methods are only applicable in special cases, such as cone beam computed tomography (CBCT) to CT registration [81], where the HUs are similar, due to the highly similar underlying imaging physics. Otherwise, feature-based approaches are applied. To extract and successfully match features or landmarks between the two involved modalities, the same anatomical structures have to be present in both.

3.1.3 Rigid and Nonrigid Registration

The field of registration can also clearly be separated into the subfields of rigid and nonrigid registration techniques. In rigid registration the transformation can be described by a rotation, a translation, and occasionally a scaling. In nonrigid registration, one of the shapes to register is deformed to have a high correspondence with the shape it is being registered to. Nonrigid registration is often referred to as elastic registration. Such methods are, e.g., thin plate

spline-based [82], free-form deformation [83], or diffeomorphic demons [84] approaches. A review article of geometric transformations for nonrigid registration was published in 2008 [85].

3.1.4 Dimensionality – 3D/2D Registration

Approaches can also be categorized based on the dimensionality of the registration problem. There are algorithms that operate on images of the same dimensions, such as methods for 2D/2D registration of X-ray images and scintigraphy [86], or 3D/3D registration of MR to CT data [87]. However, especially in image guidance, the datasets to register have predominantly different dimensionalities. The most frequent case is registering a preoperative 3D dataset to interventional 2D data. This can be the case in X-ray or ultrasound (US) guided procedures, where preoperative data, such as CT or MR, provides additional, advantegous information. Since the current registration problem of the left ventricle (LV) for CRT is also a 3D/2D problem, the focus in this chapter is on 3D/2D registration approaches.

A comprehensive review of 3D/2D registration methods was published recently [88]. The authors define categories for 3D/2D registration algorithms 1) based on how the dimensionality problem is resolved and 2) what property of the image data is used for the registration. The three categories of dimensionality are: 1) methods based on projection from 3D into the 2D image plane, 2) methods based on backprojection from the 2D plane into 3D, and 3) methods based on 3D reconstruction from multiple 2D views. The information, based on what the registration is performed, is categorized into 1) features, 2) intensity, and 3) gradients, see Figure 3.1. In this review, gradient-based methods are not handled as a separate category, since gradients can be considered as features, computed from the intensity values of an image.

The paper does not focus on cardiac image registration in particular and does only list a few methods that can be relevant for interventional guidance of cardiac procedures, such as [89] or [90].

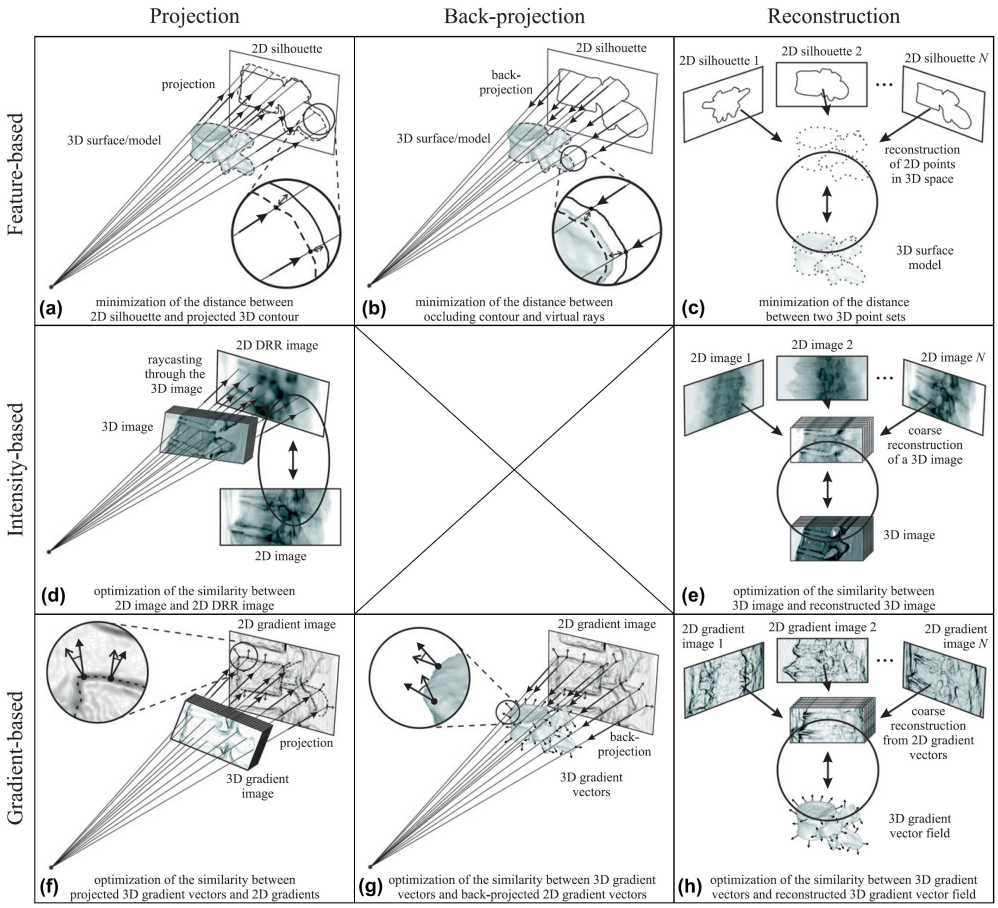


Figure 3.1: Main subcategories of 3D/2D registration in [88]. Reprinted from Medical Image Analysis, 16.3, P. Markelj, D. Tomaževič, B. Likar, and F. Pernuš, A Review of 3D/2D Registration Methods for Image-Guided Interventions, pp. 642-661, Copyright (2012), with permission from Elsevier.

3.1.5 Learning-based Registration

In recent years, machine learning (ML) approaches have proven that they can provide the performance of, or even outperform, classical approaches in multiple applications of computer vision, such as the ImageNet classification [91], semantic segmentation [92], or pose estimation [93]. After the great success in computer vision, the medical imaging community has started to adopt and develop learning-based approaches.

ML methods are widely applied for medical image segmentation [94, 95] and have started to emerge in medical image registration as well. There are different ways to involve ML in the registration workflow. One popular approach is to perform feature-based registration of features that are extracted from the datasets to be registered through ML [96, 97]. ML methods, such as random forests, can be used to synthesize images of different modalities [98] to simplify multimodality registration. A further promising approach is to learn a regression model to update the registration during optimization [99, 100].

Modern approaches based on neural networks (NNs) or deep learning have shown promising results too. A method based on a convolutional neural network (CNN) is able to regress the registration parameters of surgical implants between 3D CT volumes and 2D X-ray images [101], although with a limited capture range. A further approach is capable of iteratively predicting the transformation parameters [102] through deep reinforcement learning [103] for 3D/3D registration problems between CT and CBCT or MR and CT. Furthermore, the concept was proven to perform well in 3D/2D registration problems [55]. The approach was also extended to nonrigid registration problems [104].

Although ML approaches show promising results, approaches involving NNs are especially data hungry. The amount of data required to train a system is in the range of thousands of datasets. To train a NN, not just the images to register, but the ground truth alignment for each image pair is required. In medical image registration, this amount of data with ground truth alignments is rarely feasible to acquire. If a system is trained only on a small number of

examples, the learned model will overfit to the training samples. To overcome the problem of overfitting, data augmentation can be performed [91]. The data can be transformed by rigid or nonrigid transformations (depending on the application) to increase the number of training samples, thus to avoid overfitting. An additional way of avoiding overfitting is introducing dropout into the training process [105].

3.2 Registration for Cardiac Interventions

Registration of datasets acquired with different modalities can facilitate the representation of complementary anatomical and physiological information in one common coordinate frame. Thus, it can support the planning and guidance of interventional procedures. Minimally invasive cardiac interventions, especially electrophysiology (EP) procedures, are often conducted under X-ray fluoroscopy guidance. X-ray fluoroscopy is an ideal modality for highlighting radiopaque structures, i.e., bones and catheters, however, it does lack soft tissue contrast. To visualize soft tissue during the procedure, such as the chambers of the heart or its main vessels, a 3D/2D overlay can be provided through an accurate registration of a preoperatively acquired dataset, such as (segmented) CT or MR data.

Registration of cardiac structures for image guidance of minimally invasive interventions is a highly challenging task. The heart is a constantly, periodically moving organ that has no tightly fixed location in the thorax related to other organs. Its periodic motion is a combination of respiratory and cardiac motion and its relative pose might change due to patient movement. This becomes relevant if the patient is in a different pose during the intervention than during the preoperative scan, the patient moves, or is being moved during the intervention. Due to the possible movement of the heart, registration methods using surrounding structures, such as the spine or the ribs are not reliable [5].

An even more challenging problem is that the heart barely has any landmarks, in some cases even none, that are stable and well visible in multiple

modalities. Particularly in the case of this work, cine MR to X-ray fluoroscopy registration, there are no easily identifiable, strong, shared landmarks visible.

An extensive review of registration methods for cardiac procedures was published in 2002 [106]. The article considers mainly 3D modalities in order to perform 3D/3D registration for preoperative planning. The modalities included are CT, magnetic resonance imaging (MRI), positron emission tomography (PET), single photon emission computed tomography (SPECT), and, in a few registration methods, echocardiography. The review article describes 3D/3D registrations with these modalities well, however, does not consider 2D modalities, such as X-ray fluoroscopy, thus no 3D/2D registration is studied in depth. This suggests that the methods described cannot be directly applied to register images for intraoperative guidance. A more recent review article considers 2D modalities and 3D/2D cardiac registration methods briefly [107]. In the following sections, an extensive review of registration methods, that are relevant or were explicitly developed for CRT procedures, is given.

3.2.1 Registration for Guidance of CRT Delivery

A brief enumeration of existing methods for guidance of CRT procedures by 3D/2D image fusion with X-ray fluoroscopy was published in 2011 [108]. The article handles methods for registering MR or CT data by manually aligning previously extracted 3D models to identifiable objects in X-ray fluoroscopy, such as the heart shadow and catheters. It also considers automatic registration by using the spine, reconstructed catheters, or an MR/X-ray suite (XMR) system (a combination of an interventional X-ray and an MR system). The article states that 3D US registration is challenging and it is usually performed indirectly through registering the US data to a CT or an MR volume in the first instance, or directly by registering through a tracked US probe. The work mentions that the registration of CBCT to X-ray fluoroscopy is implicit during a procedure in the catheterization laboratory, since the datasets are recorded with the same device in the same coordinate frame.

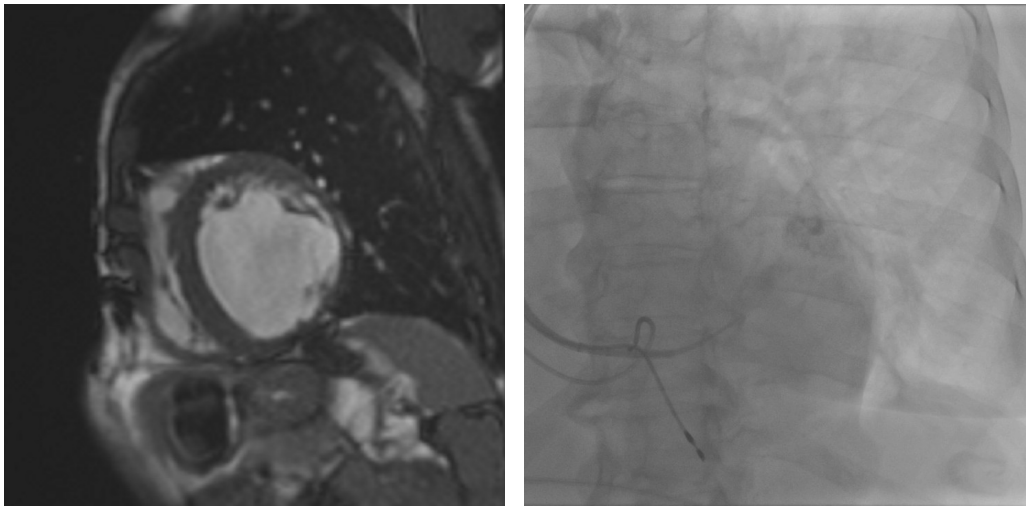
The CRT related registration literature is categorized into 1) manual, 2) external marker- or tracking-based, 3) endovascular device-based, and 4) anatomical structure-based methods. The challenges of registering MR to X-ray data and methods (those presented in [108] and more recent approaches) that can be considered for 3D/2D registration for CRT delivery guidance will be described in more detail in the following sections.

3.2.1.1 The Challenge of Cardiac MR to X-ray Registration

As mentioned in Section 3.2, the main challenge in registering intermodality data is the lack of shared information between the modalities to register. This is especially exacerbated in MR to X-ray registration. The main reason is the fundamentally different physics of the image formation. MRI can image soft tissue with high water content, due to the bound hydrogen in the water molecules. X-ray, on the other hand, excels at imaging hard, radiologically dense (radiopaque) tissue types, such as bones. The object to register, the heart, mainly consists of muscle tissue of high water content, but its radiopacity is low. Thus, it is well represented in MR images, but only a light shadow is visible in X-ray images, see Figure 3.2.

The second main complication, the difference in dimensionality, is highlighted in Section 3.1.4. The cine MR images of standard cardiac protocols are acquired slicewise, resulting in a 3D volume if compiled. This discrepancy of image dimensionality can be resolved through projection from 3D to 2D, reprojecting or reconstructing to 3D from 2D acquisitions, as described in Section 3.1.4. Further complications are resulting from different fields of view (FOVs) and resolutions. Especially the out-of-plane resolution of standard short axis (SAx) MR images can be extremely low (approx. 10 mm) compared to the submillimeter resolution of X-ray images.

Due to the complications described above, it is extremely challenging to develop automatic registration approaches for cardiac MRI to X-ray data. Thus, most current frameworks for image guidance rely on manual, human interaction.



(a) SAx MR (b) X-ray

Figure 3.2: Difference in appearence of the heart in MR and X-ray images. (a) Cine MR acquisition frame of a SAx slice. (b) Interventional X-ray acquisition frame.

3.2.1.2 Manual Methods

The most basic approach of aligning images is to perform the registration manually. A clinical expert can identify the landmarks in two datasets to register and define the necessary transformation by hand. For CRT, a loop catheter can be used for registration [9]. The idea is to position the catheter inside the right atrium of the heart in order to reduce the difficulty of manual registration from two angulations. The method was evaluated in seven clinical cases. The target registration error (TRE) was measured in 2D between the centerline of the coronary sinus (CS) from the MR and from contrasted X-ray fluoroscopy. The 2D TRE was found to be 1.30 ± 0.68 mm. However, the measurements were conducted between ten points along one of the centerlines and the ten nearest points of the second centerline. Thus, the 2D TRE does not just exclude a measurement in depth, but also does not incorporate a shift along the centerlines either. Furthermore, the measure does not account for the foreshortening of the vessel in the X-ray images, that would introduce another factor to the 3D TRE. It is to be noted that a 3D whole heart MRI was acquired to visualize the left atrium (LA) for registration, and the CS for evaluation. Classical 3D MRI is not feasible in clinical practice, due to the difficulty of lengthy breathholds (20 seconds or

longer) for heart failure (HF) patients. Even the standard cardiac MR protocol can be exhausting with breathholds of ca. 10 seconds. Free-breathing 3D MRI sequences exist [109], however, they are not widely available. The majority of clinical workflows rely on 2D cine MRI acquisitions.

Manual approaches work in those scenarios very well where the two modalities incorporate a sufficient amount of shared information, common landmarks, such as the coronary venous anatomy. Without easily identifiable landmarks, the registration process can be time-consuming, inaccurate, and operator-dependent. In such scenarios, automatic approaches are preferred.

3.2.1.3 External Marker- and Tracking-based Methods

In registering X-ray images with a 3D modality for cardiac interventions there are only a few cross-modality landmarks available. Early approaches tried to solve the registration problem by applying external markers to the patient's body, or the imaging system. One of the first means to approach the problem this way was developed in 2003 [89]. The developed system is among the first hybrid X-ray and MR systems, also known as XMR systems. The MR and the C-arm X-ray systems are installed in the same room and calibrated by phantom measurements to find the transformation between the two systems. For the calibration and interventional registration, infrared light emitting diodes (IREDs) and an optical tracking system are used, see Figure 3.3.

The registration consists of a chain of steps. At first, the relation of the X-ray table and the C-arm system is determined over an optical tracking system. Secondly, the C-arm system is calibrated by imaging a calibration phantom from multiple angulations. The relation of the MR image space and the MR table space is implicitly stored in the recorded image headers by the system. Afterwards, the transformation between C-arm space and the MR space is calculated. A calibration phantom is imaged and its manually annotated markers are compared with the markers in C-arm space, acquired in 3D by the optical tracking device. Finally, the translation of the sliding tabletop is determined by a combination

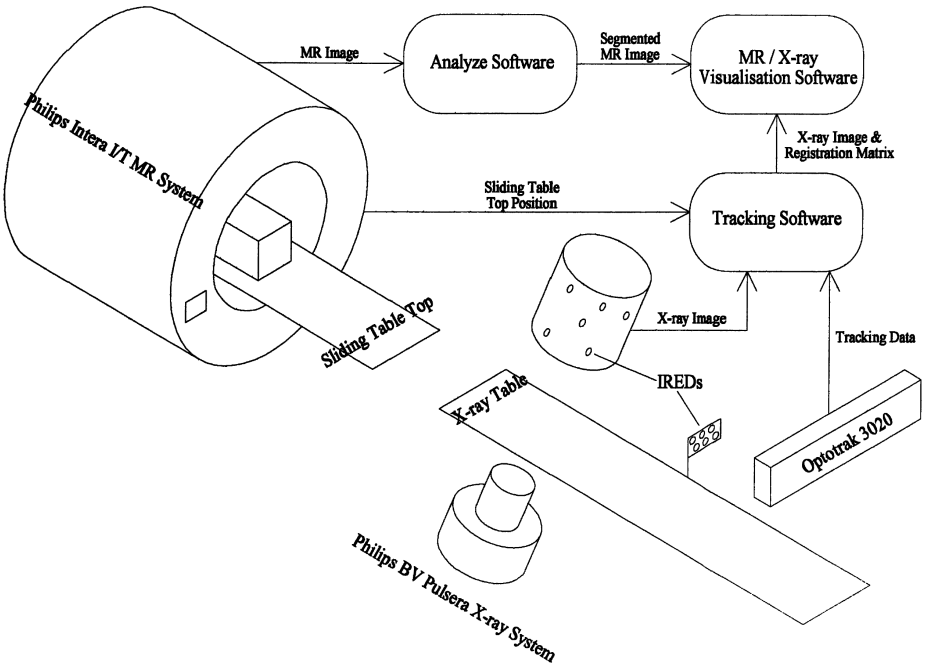


Figure 3.3: XMR setup and registration framework with internal table and optical tracking [110]. © 2005 IEEE.

of internal MR tracking and the optical tracking system. The described steps and the interventional projection matrices of the C-arm, result in an accurate registration between the MR and the C-arm X-ray system.

The evaluation of the accuracy was performed in phantom studies. A point-based phantom was registered with a 2D TRE of 4.2 mm and a 3D TRE of 4.6 mm. For an antropomorphic vessel phantom the 2D TRE was 3.6 mm and the 3D TRE was 5.1 mm. Two clinical cases were also presented. Two XMR guided cases registered by the registration method are described in more detail in [111]. Eleven successfully conducted cases were reported in [110] with this approach. The method was developed for minimally invasive cardiac procedures in general and it can be applied in CRT interventions.

The method performs well for a rigid setup (phantom experiments), but no quantitative evaluation on patient data was presented. Evaluation in patient data is important, since patient movement could negatively affect the registration, thus could greatly diminish results. Even if the patient is transferred directly to the X-ray system after the MR acquisition, the body might move, or might be moved

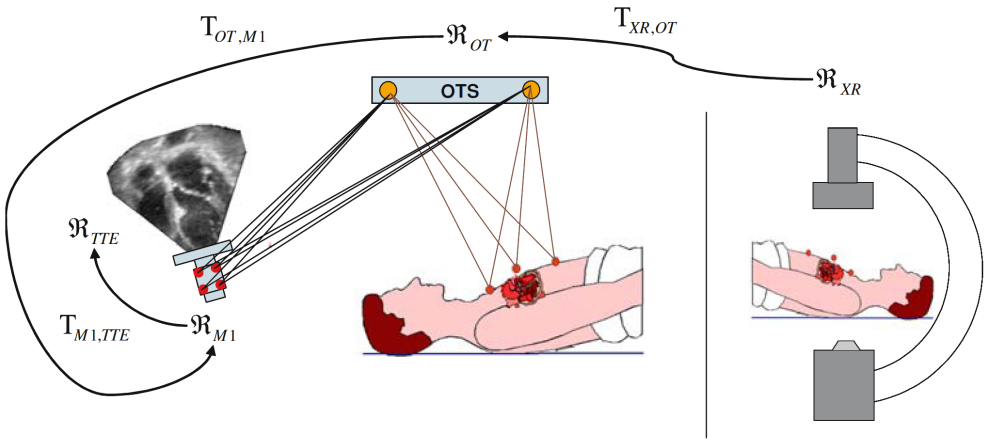


Figure 3.4: US to C-arm registration setup showing the step by step transformations [112]. Reprinted by permission from Springer Nature: Springer Nature International Journal of Computer Assisted Radiology and Surgery, Fusion of three-dimensional X-ray angiography and three-dimensional echocardiography, Volker Rasche, Moussa Mansour, Vivek Reddy et al, © (2008).

accidentally. Furthermore, relative displacement of internal organs, especially of the heart, is not compensated for. Additionally, even if the system performs adequately, the availability of XMR systems is highly limited and installing additional equipment in the catheterization laboratory might not be feasible.

A similar approach was presented for the registration of 3D US data to 3D rotational X-ray angiography (RXA) [112], see Figure 3.4. External fiducial markers are attached to the patient's chest. Although the markers are well visible in X-ray images, they cannot be imaged by the transthoracic echocardiography (TTE) probe. The missing transformations are recovered, similarly to [89], by utilizing an optical tracking system. First, the relation of the TTE probe and the TTE image is determined by a phantom experiment. Fiducial markers are applied to a phantom that are visible by the TTE probe and markers that are visible by the optical tracking system. The relation of the optical and the US markers is known. Trackable markers are applied to the TTE probe as well, thus the optical tracking system can easily determine the transformation from image space to TTE probe space. After the transformation is known, since the TTE probe and the chest markers are both optically tracked, the transformation from the TTE image space to the optical tracker space is known.

During the intervention, the chest markers appear on the 3D RXA images and can be easily aligned with the chest markers in the optical tracker space by calculating the singular value decomposition (SVD). The resulting registration is used for fusion of anatomical data from the RXA image (coronary veins) and functional data (mechanical activation, strain) from the TTE images.

The method was evaluated quantitatively, similarly to the XMR fusion, only in phantom data and first experiments were made with real patients. The 3D TRE in phantoms was below 2 mm. The position of the coronary veins appeared to be reasonable after the registration was performed during live clinical cases. A quantitative evaluation was not performed on patient data. The system is prone to errors induced by patient motion, but it is more robust than the XMR setup. The patient may be in different positions during the two scans, however, since the fiducial markers are attached to the chest, the aligment can still be performed. If the heart's relative position to the fiducials changes, the registration will become invalid. Furthermore, the relative position of the fiducial markers to each other may change too, due to the elasticity of the skin surface.

A simpler method was published for MR to X-ray registration [113]. External, multimodal, fiducial markers are attached to the patient's chest before the MR scan. The patient is transferred into the catheterization laboratory and imaged with the fiducial markers still applied. Since the markers are well visible in both modalities, the points are marked and registered by an iterative nonlinear least squares optimization method.

The method was evaluated in vivo in a porcine model, performing endomyocardial injections. The locations were defined in the myocardium before the injections and a specific MR protocol was used that could highlight the injected agent after the procedure. The preinterventionally defined positions were compared with the post-injection measurements. The measurements show a 3D TRE of 3.2 ± 2.6 mm.

The described registration methods are able to register the heart accurately, however, only under special circumstances. The biggest restriction being that the

attached markers have to stay on the patient's body from the preoperative scan to the interventional imaging. If the preoperative images are acquired weeks or even months before the intervention, the approach is not feasible. This is obviously not a problem for the XMR registration [89], since the intraoperative imaging happens directly after the preoperative scan. A further major potential problem is that, even if the intervention is in immediate succession of the MR scan, the heart may change position and orientation by the time of the intervention, thus invalidating the registration.

3.2.1.4 Endovascular Device-based Methods

Registration for CRT delivery guidance can be performed by involving endovascular devices. Endovascular devices are catheters that are inserted through vascular access. They can be associated with an anatomical structure, such as a coronary vessel or a chamber of the heart, in order to facilitate registration between them.

There are several approaches published that use a 3D preoperative modality to extract the medial line of a vessel, i.e. the CS, in order to register it to a catheter reconstructed from fluoroscopy during the intervention [114, 6, 7]. The method described and evaluated in [114, 6] is designed for cardiac interventions in general. A global search is performed to find the correct registration. The method was qualitatively evaluated in two catheterization cases, of which only one was successful [6]. In [114] the qualitative evaluation was successful in two of three cases. No quantitative evaluation was performed. It was noted, that the algorithm provides better results, if not only one, but multiple mapping catheters are being used under some conditions. However, the computation time increases drastically, making the application of the method unfeasible in an interventional setting.

A highly similar approach was developed for atrial fibrillation [7]. Instead of the extraction of the whole CS centerline, only the starting point and the end point (where the CS continues as the great cardiac vein) are identified in a whole

heart MR sequence. These points are registered by an iterative closest point (ICP) algorithm to the reconstructed catheter in the CS during the intervention. This method was evaluated quantitatively on in vivo data of eight patients. The average TRE between the automatic and a manual registration by an expert was 5.7 mm.

A method was also developed that segments the great vessels in the close proximity of the heart from CT data to register to inserted catheters [115]. The approach was extended by registering with multiple catheters and applying a weighting to the registration of a vessel to a catheter. The applied weight depends on the diameter of the vessel being registered. The higher the diameter, the lower the weight is, since the probability of the catheter being close to the centerline of the vessel decreases. The algorithm was evaluated with several combinations of great vessels in a heart phantom. The 3D TRE decreased in almost all of the combinations. It was 0.55 mm if considering anatomical landmarks associated with all four chambers of the heart.

The endovascular device-based methods described by now show high similarity. Recently, a novel approach was published for catheter-based registration [8]. In this method, a catheter is simulated in the segmented volume of the CS from a 3D whole heart MR volume. The simulation was performed with a finite elements method and the stiffness of the catheter was measured beforehand. During the intervention, first, an initial, manual alignment is performed. Then the simulated catheter is projected into the 2D plane of the X-ray image. A gradient field is computed for the projection and the X-ray image to set up a NCC-based metric between the two. The NCC-based metric is finally heuristically optimized to find the registration matrix. The algorithm was evaluated in phantom experiments. The accuracy was reported to be sufficient for CRT procedures, since the 3D TRE was below 6 mm, the reported approximate average half-size of a 17-segment American Heart Association (AHA) bull's eye plot (BEP) segment [8]. The method distinguishes itself from the previous approaches, since it uses a single X-ray image and not biplane images for the registration. This, however,

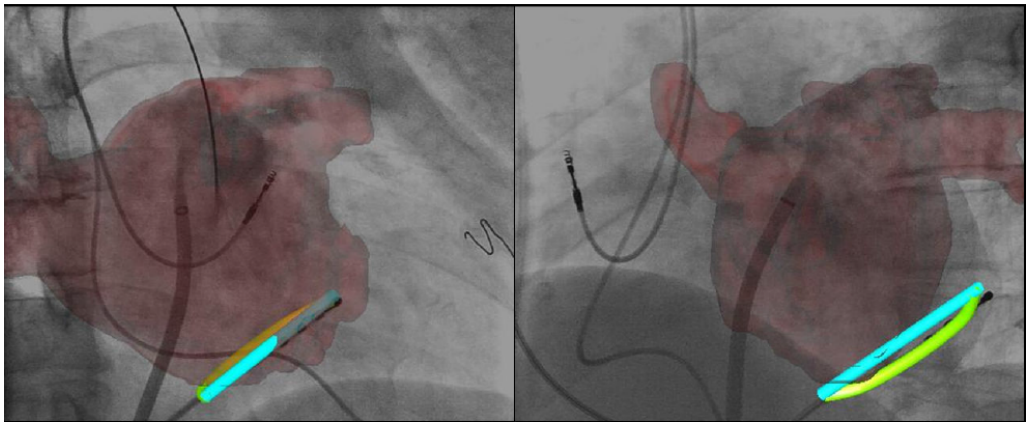


Figure 3.5: Interventional overlay of preoperative information (left atrium in red). The preoperative CS centerline (CL) (green) is not congruent with the intraoperative CL of the catheter (blue) in the CS [7].

means that the through-plane translation cannot be estimated accurately. If the through-plane alignment is critical, a second angulation can be involved into the registration.

Most of the described endovascular device-based methods show sufficiently low errors for CRT guidance, however, most of the measurements were done in rigid phantoms. During a CRT case, the inserted devices could deform the anatomy of the vessels, thus might invalidate the preoperative models, see Figure 3.5. This behaviour would result in higher errors that might be above a tolerable level for a CRT intervention.

3.2.1.5 Anatomical Structure-based Methods

The third group of registration methods for CRT guidance is based on anatomical structures, such as coronary vessels or grooves of the heart. The use of internal structures has the benefit that inserted devices, such as the catheters in the endovascular device-based methods, are not present or their presence is limited, thus the anatomy is not distorted significantly.

The simplest approaches utilize the cardiac vessels as landmarks for the registration. In [90] a vessel-to-vessel method is presented between a segmented coronary artery mesh model and contrasted X-ray fluoroscopy images. The coronary mesh model could be extracted from any 3D modality, e.g., CT or MR.

The model is projected into segmented, mono- or biplane, contrasted fluoroscopy images and a similarity measure based on the entropy correlation coefficient (ECC) is optimized. High accuracy was reported: 1.60 ± 0.21 mm in mono- and 0.53 ± 0.08 mm in biplane registration. However, solely virtual experiments were performed. In patients, the method might not find an adequate solution, due to numerous local minima being potentially present.

The idea was extended by introducing a nonrigid approach for vessel-to-vessel registration [116]. The algorithm performs the rigid alignment similarly to [90], but initiates a nonrigid refinement afterwards. The nonrigid adjustment is intended to be useful for compensating for cardiac and respiratory motion induced deformations in the anatomy of the vessels.

Other landmark-based approaches are using the grooves and the surface of one chamber of the heart, i.e., the LV [10, 11]. It was proposed to segment SPECT images of the LV to extract a model of the myocardium. Due to the low resolution of the SPECT data (approx. 10 mm $\times$ 10 mm pixels), the wall thickness was defined to be constant, 1 cm. The interventricular grooves were defined based on the intensity variations of the SPECT images. During the intervention, the vascular tree of a coronary artery is segmented from contrasted X-ray fluoroscopy. The coronary tree is matched with the corresponding groove automatically to provide an initial alignment for a local ICP refinement between the vessels and the LV surface. The rigidly registered vessel models are warped onto the surface of the LV.

The approach described in [11] is highly similar and it was developed specifically for CRT. The method operates on SPECT images and two angiograms of the coronary veins. The alignment is initialized by geometry information from the dicom headers. The least squares solution of the vessels and the corresponding grooves is computed. The rigidly registered venous tree is mapped onto the surface of the ventricle similarly to that in [10]. In [11] a more extensive evaluation is provided. The coronary vessels are extracted from a CT volume and are registered with the SPECT data. The position of the reconstructed vessels is

compared against the vessels from the CT volume. The accuracy is stated to be sufficient for CRT guidance, since all the errors were within the segment size of the 17-segment AHA plot (30 mm $\times$ 30 mm) [11].

These methods have several assumptions that are questionable. Due to the low resolution of the SPECT data (ca. 10 mm pixel spacing [117]), the LV model and the defined groove landmarks can be inaccurate. The assumption that the vessels lie in the grooves of the heart is not always valid, due to the high variability of venous anatomy [14]. The final warping of the vessels does also not promote an accurate registration for interventional guidance, rather a visually appealing representation of the fused datasets, e.g., for preprocedural planning.

Cardiac registration was also studied for CRT between electro-anatomical mapping (EAM) data and other 3D datasets, such as MR, CT, or echocardiography [118, 119, 120] and for 3D+t MR to 3D+t CT data [87]. These methods of image fusion were designed for preoperative planning and are not likely to be applicable for interventional guidance.

3.3 Evaluation of Manual Registration

The above methods are designed for cardiac interventions, but they are not utilized in current clinical practice. If a preoperative image is used for image guidance, manual registration still represents the state-of-the-art. To compare with the currently established clinical method, thus to motivate this work, a manual registration experiment was performed. Unfortunately, there is no dataset available of paired MR and X-ray data with known ground truth (GT) transformations. The hypothesis is that accurate manual registration of this kind of data is not reliably possible, thus we cannot generate ground truth through manual registration between MR and X-ray data.

To have data with ground truth registrations, CT data is used instead of MRI to extract the mesh models. This has the benefit that X-ray-similar CT projection images, also known as digitally reconstructed radiographs (DRRs),

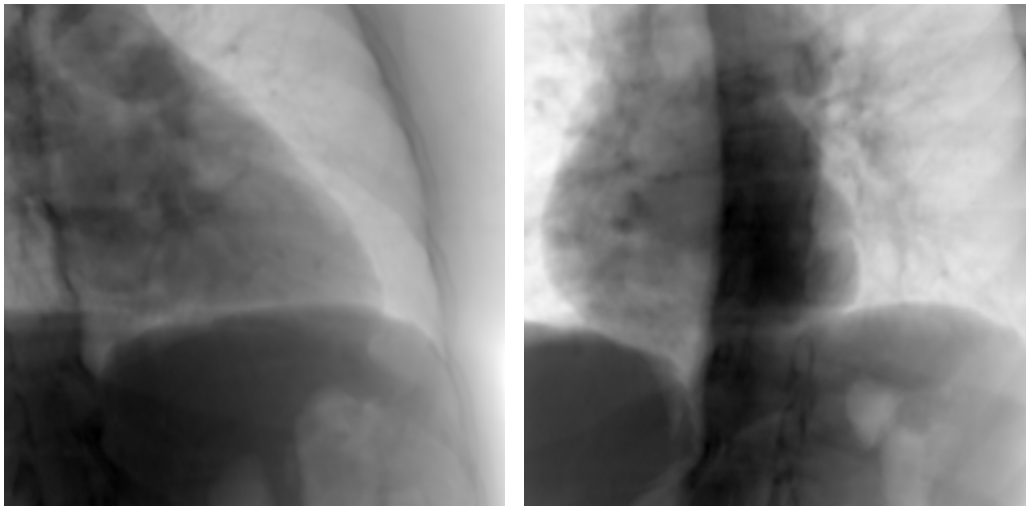
can be generated. The DRRs are generated by a predefined projection geometry, thus the ground truth registration (with no error) is known.

CT data of ten patients is used to generate two projection images each, right anterior oblique (RAO) 30° and left anterior oblique (LAO) 30°, representing typical X-ray projections used for manual registration in CRT procedures, see Figure 3.6. The epicardial LV mesh models were extracted through automatic segmentation [121]. The models were perturbed with a randomly generated 3D transformation of a uniform distribution of six degrees of freedom (DOF), three translations and three rotations. The perturbation was in the range of ± 35 mm along each axis of translation and $\pm 15°$ about each axis. The ranges are representative of misalignment observed in interventions before registration. This process resulted in ten cases in total.

The registration was performed by seven non-clinical experts. The ten cases were registered by the experts on average in 35 min. The results show that the mesh (vertex-to-vertex) mean absolute error (MAE) was 12.03 ± 4.91 mm of all cases. The accuracy is rather low, but it can be sufficient for conventional CRT interventions, however, the large variation of the TRE shows, that the registration results are highly user-dependent. In seven cases, the difference between the minimum and maximum error was larger than 10 mm. These findings suggest, that an automatic approach would be beneficial that could eliminate the user-dependency of the registration results, thus provide a more reliable overlay for image guidance.

3.4 Discussion and Conclusion

A substantial amount of progress was made in the field of medical image registration for interventional guidance in the recent years. There have been several methods developed for cardiac interventions that were applied or could be applied to CRT interventions, however, most of them use a truly 3D modality, such as CT or 3D (whole heart) MR to acquire preoperative data. Since CT means additional dose to the patient and currently cannot provide the necessary scar



(a) RAO DRR (b) LAO DRR

Figure 3.6: DRRs generated from a CT volume to assess manual registration performance. (a) RAO 30° and (b) LAO 30° projections to resemble standard CRT working angulations.

tissue imaging, MR is preferred for CRT. Classical 3D MR images are challenging for heart failure patients, because several long breath-holds are necessary for an acquisition and the scan may take more than 30 minutes, thus can be exhausting for patients. Due to novel free-breathing 3D MRI being not widely available, it was proposed to use 2D cine MR acquisitions for CRT procedure planning and guidance, however, the registration between X-rays and cine MR images is an open research problem. It poses a major challenge that only a few, or no common landmarks, are present.

Due to the lack of landmarks, manual registration is challenging, requires a high level of experience from the operator. Optical tracking-based approaches are prone to patient motion caused misalignment, and just as fiducial marker-based alignment, cannot compensate for the motion of the heart relative to the skin surface. Inserted endovascular devices can distort the anatomy, thus invalidate the preoperative models and 3D whole heart MR data is not available to extract the coronary veins. Approaches performing the registration based purely on the anatomy of coronary vessels of the LV are not feasible due to the lack of the 3D MR data to extract the vascular anatomy. Previous approaches to register

intraoperatively extracted vessel models and preoperative models of the LV have several assumptions, such as that some of the vessels lay in specific grooves of the heart, i.e., the interventricular or the atrioventricular groove. These assumptions are questionable and the extraction of the preoperative models and grooves can be inaccurate, due to the resolution of the 3D modality used, i.e. SPECT.

After the above analysis of previous methods, a novel approach for cardiac image registration is proposed: registration by adjacent anatomical structures. A similar approach was proposed for cortical data from CT and MR, where the skull is registered to the surface of the brain, or a specific region of the brain to blood vessels embracing it [122]. The approach does not need cross-modality landmarks. The registration is performed by exploiting prior knowledge about the adjacency of anatomical structures. The adjacent structures are the LV extracted from MR and the coronary veins reconstructed from contrasted, interventional X-ray images. The venous network is represented as a point cloud and the LV as a surface mesh model. A point cloud matching algorithm, such as ICP [123], coherent point drift (CPD) [124], or a globally optimal variant: globally optimal iterative closest point (GO-ICP) [125, 126], can be used to register the vascular point cloud to the surface of the LV model. This approach represents a novel, automatic way of registering cardiac data without cross-modality information, by aligning adjacent anatomical models. The adjacent anatomical model-based registration approach is described in the next chapter in more detail.

Chapter 4

Registration through Adjacent Anatomical Models

The previous chapter has described cardiac registration approaches for interventional guidance, outlining their technical and clinical limitations. This chapter presents an approach that can register mesh models extracted from preoperative magnetic resonance (MR) to a vessel point cloud reconstructed from contrasted, interventional X-ray images. A thorough qualitative and quantitative evaluation of the developed approach is performed on phantom and clinically acquired cardiac resynchronization therapy (CRT) patient datasets.

4.1 Introduction

Chapter 2 introduced that patients with advanced drug-refractory heart failure can be safely treated with CRT. However, 30 to 50 % of patients do not respond to therapy [2]. In this procedure, a CRT device, a biventricular pacemaker, is implanted using fluoroscopic image guidance. Suboptimal placement of the lead on the left ventricle (LV) has been identified as a leading cause of non-response. Unfortunately, improving the placement of this lead is extremely challenging for clinicians.

It has been shown that placing the LV lead away from scar tissue and in the latest point of mechanical activation can improve the response rate [2]. This information cannot be directly obtained from fluoroscopic images and requires an

additional imaging modality, such as preoperative magnetic resonance imaging (MRI). Registering preoperative MRI with intraoperative fluoroscopic images enables clinicians to visualize scar and mechanical activation in real time fluoroscopic overlay, guiding the placement of the LV lead to improve response rates.

It was highlighted in Chapter 3 that registering non contrast-enhanced MRI to X-ray fluoroscopy remains an open research problem. The main challenge is the lack of shared information between the modalities. In the case of cardiac MRI, the images show high soft tissue contrast to visualize the heart, but bony structures, such as the spine, are not easily seen. In fluoroscopic images, however, bony structures and instruments are well visible, but there is a lack of contrast for soft tissue. A way to enhance the visualization of soft tissue anatomy is to inject contrast agent. In CRT this is used to visualize the coronary sinus (CS). The registration problem is further complicated by different fields of view (FOVs), the low resolution of MRI, and cardiac and respiratory motions.

Current approaches for registering MRI to fluoroscopy can be categorized as manual [9], fiducial- [113, 89], tool- [127, 6, 8], or anatomical landmark-based methods [10, 11]. The above approaches are mostly not suited to standard clinical workflows. In current clinical practice, manual registration is still the dominant method. It is technically simple, but practically challenging, due to the nature of registering a 3D image to a 2D image(s) without strong landmarks. The process can be time-consuming, inaccurate, and highly user-dependent, as shown in Section 3.3.

The goal is to develop a registration approach that 1) has a workflow suitable to all hospitals, e.g., does not require fiducials, 2) does not require lengthy additional MR acquisitions, and 3) can be used early in the procedure before tools are inserted into vessels.

This chapter presents a method for registering adjacent anatomical structures, similarly to [122], but for cardiac registration, and introduces the concepts of superabundant vessel reconstruction and dynamic outlier rejection. For CRT

procedures, the adjacent anatomical structures to register are the LV (from non contrast-enhanced cine MRI) and the coronary veins (from contrast-enhanced X-ray fluoroscopy). This approach is highly suitable for X-ray to MRI registration as it does not use cross-modality information. However, it requires the vessel system to be reconstructed from two X-ray images and the relatively sparse vessel system to be registered to the surface of the LV.

Reconstructing the vessel system is challenging in CRT where contrast agent washes out of veins quickly and the two X-ray images are acquired sequentially with separate and potentially inconsistent contrast injections. The problem is exacerbated by varying quality of CS ostium occlusion and missing vessels. For reconstructing coronary arteries, approaches that match bifurcations have been proposed [128]. However, for CRT, coronary vein reconstruction techniques require manual vessel matching [10, 11]. This work proposes superabundant vessel reconstruction, in that all possible epipolar correspondences are reconstructed, which explicitly removes the need to perform vessel matching. The reconstructed vessel model is guaranteed to contain the true vessel structure and additional outliers.

The reconstructed coronary vein model is sparse relative to the LV surface. Registering these two structures is a partial surface registration problem which is well known to get stuck in local minima, especially when the data contains outliers. To overcome this challenge, a novel dynamic outlier rejection method is proposed. Outliers are estimated in the vessel reconstruction phase and combined with a globally optimal registration framework to prevent the registration from becoming stuck in local minima.

The accuracy of the proposed system is evaluated on phantom and in vivo data. A novel validation framework is proposed that exploits artificial valves implanted in the patients. The clinical application is demonstrated on patient data and evaluated by surveying a team of clinical experts.

4.2 Methods

In the following sections, the concepts for adjacent anatomy-based registration by superabundant reconstruction are described. First, a short overview is given about the methods involved, then the extraction of the models to register and the registration approach are described in more detail.

4.2.1 Overview

Classical registration approaches for medical images from different modalities rely on matching anatomical structures or landmarks that are present and visible in both images. These are known as cross-modality landmarks. MR and X-ray fluoroscopy images are intrinsically different and cross-modality landmarks are often not easily identifiable or even present. This can be explained by the inherently different imaging physics of the two modalities. This work proposes a novel approach that uses anatomical structures that are adjacent to each other. The two adjacent structures in CRT are: 1) the LV epicardium (from MRI) and 2) the CS and its tributary coronary veins (from fluoroscopy), the venous anatomy on the surface of the LV.

The main steps of the proposed method are depicted in Figure 4.1. Pre-operative MRI images are automatically segmented and a 3D model of the LV epicardium is generated. Intraoperatively, the coronary veins are extracted from two frame-matched fluoroscopic images and a superabundant 3D vessel model is automatically reconstructed. Finally, the 3D vessel model is registered to the LV epicardial mesh with a globally optimal iterative closest point (GO-ICP) method [125, 126] where outliers are modelled using information derived from the superabundant vessel reconstruction step.

4.2.2 LV Epicardial Model

The epicardial mesh is generated from the automatically segmented MRI. The short axis (SAx) and long axis (LAx) MR images are segmented with a combination of a machine learning-based landmark detection and grey level analysis [129].

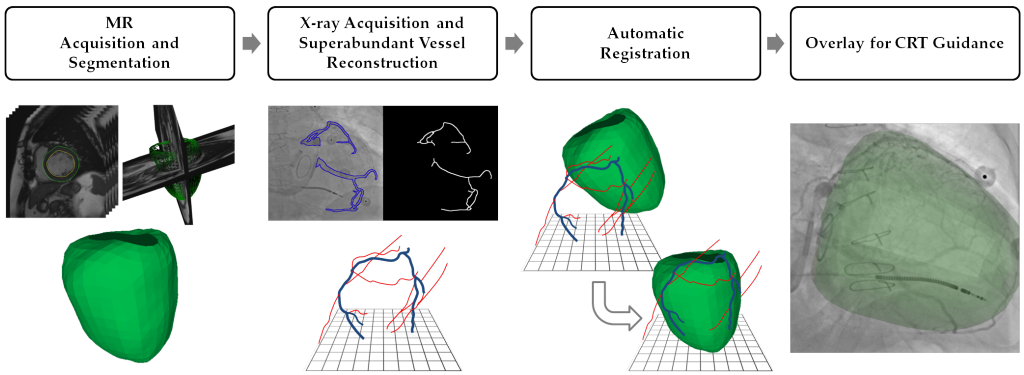
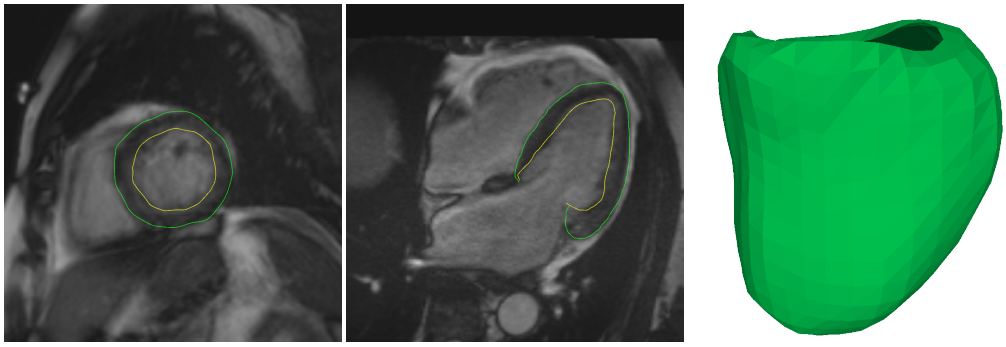


Figure 4.1: Workflow steps of proposed registration framework.



(a) SAx slice (b) 4CH slice (c) Mesh model

Figure 4.2: LV epicardial mesh model generation in 3D from segmented short axis (SAx) and four chamber (4CH) slices. The epicardial (green) and endocardial (yellow) borders are highlighted.

The contours of the epicardium are extracted using a minimum path algorithm based on histogram analysis in every slice, see Figures 4.2 (a) and 4.2 (b). The resulting contours are propagated through all phases by distortion fields and are used to generate a 3D mesh model of the epicardium for each heart phase over the cardiac cycle, see Figure 4.2 (c). The end diastolic mesh model is selected for registration, due to it being simple to identify, based on the electrocardiogram (ECG) triggering of the MR acquisition.

4.2.3 Vascular 3D Model Reconstruction

To generate the anatomical model adjacent to the LV, two X-ray images of the coronary venous tree are acquired enhanced through contrast agent injection. The vessels are segmented and the segmentation is reconstructed to receive a

superabundant, outlier-rich, point cloud.

4.2.3.1 Fluroscopic Frame Gating

In the current CRT workflow, a monoplane fluoroscopic system is used. Unlike with a biplane system, which can acquire two images simultaneously, the monoplane system must acquire the two fluoroscopic image sequences sequentially. This requires two contrast injections and two image acquisitions. The standard clinical C-arm angulations for a CRT procedure are anterior-posterior (AP) 0°, left anterior oblique (LAO) 30°, and right anterior oblique (RAO) 30°. The proposed method can use any combination of these angulations. To estimate the correct heart phase in the two acquisitions, ECG or an image-based motion gating can be performed. It is proposed to use an image-based motion gating approach that does not rely on ECG data to have a more generically applicable method. The end diastolic frame of each sequence is selected by an approach based on masked principal component analysis [130]. The method extracts cardiac motion by band pass filtering the variation of the first principal component. Furthermore, it is verified that the selected frames have a sufficient contrast agent fill for segmentation.

4.2.3.2 2D Vessel Detection

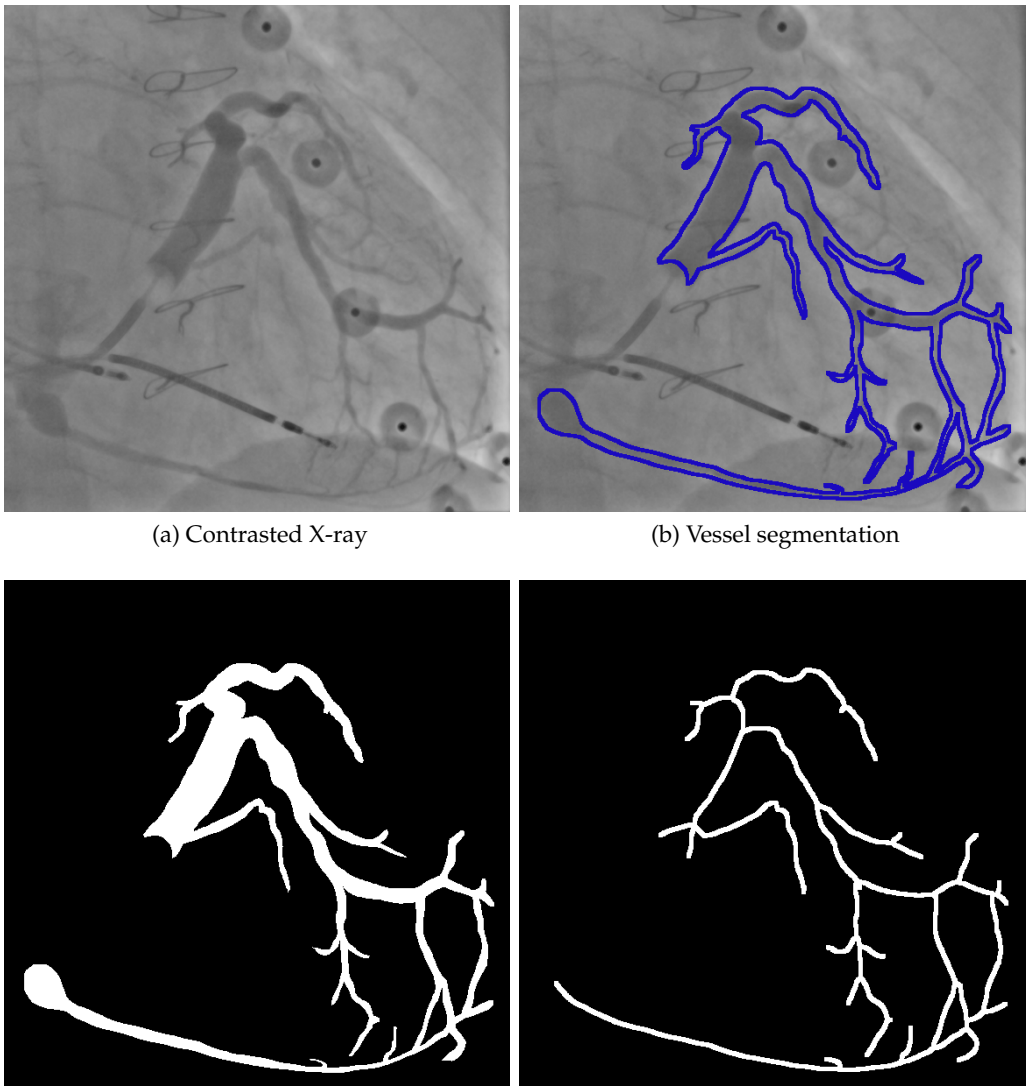
The vessels in the contrasted X-ray sequences are segmented and skeletonized to extract the centerline of the CS and its tributary coronary veins. The proposed approach is agnostic to the segmentation method, it can work with automatic, semi-automatic, or manual segmentation approaches. Automatic [131] and semi-automatic methods [132] exist for coronary vessel segmentation. In the developed framework, semi-automatic segmentation is performed. This is consistent with other registration approaches for CRT interventions [10, 11]. An average intensity image is computed from all frames of the venogram sequence. The frame to be segmented is divided by the average image of all frames of the sequence to eliminate static structures, e.g., the spine and static instruments. The resulting

image is filtered with a median filter. The Frangi vesselness filter is applied to the smooth image to enhance vascular structures [133]. The vesselness map is binarized by a manually set threshold, such that only strong response is shown. The resulting mask is manually adjusted, if necessary: misclassifications are removed, such as instruments and tubular non-vascular structures. The resulting segmentation is skeletonized [134] to create a binary mask of the centerline of the vessels. In order to efficiently and accurately process the 3D reconstruction, the binary centerline is approximated by polylines by the Ramer-Douglas-Peucker algorithm [135, 136], see Figure 4.3 for the individual steps.

4.2.3.3 Superabundant 3D Vessel Reconstruction

The extracted 2D vessels can be reconstructed in 3D if the vessel correspondence and C-arm geometry (epipolar constraint) are known. However, in CRT it is challenging to automatically detect the correct correspondences. The capture of fluoroscopy images at different angulations causes that the vessels have different visual appearance. The images are acquired with two separate manual contrast injections that may not be consistent. Contrast can be poor if the balloon is not fully deployed. Parts of the vessel system can be self-occluding or out of the field of view in one of the images and the vessel structure can be complex, causing multiple potential correspondences. Furthermore, small phase matching errors of an image pair might result in large reconstruction errors if a wrong correspondence is selected, thus could greatly diminish registration accuracy.

To address the challenges outlined above, this work proposes a fundamentally different approach for 3D reconstruction of the vessels. Instead of establishing one-to-one vessel correspondence and generating a high quality 3D model, the proposed approach uses multiple correspondences and reconstructs a superabundant 3D model. The benefit of this approach is that the reconstructed superabundant 3D model is guaranteed to contain the actual vessel structure, although it will also contain a significant amount of incorrect data or outliers.



(a) Contrasted X-ray

(b) Vessel segmentation

(c) Vessel binary mask

(d) Vessel centerlines

Figure 4.3: Vessel segmentation from contrasted X-ray images. (a) Initial contrasted X-ray frame. (b) Semi-automatic segmentation of coronary veins overlayed onto the frame. (c) Binarized segmentation image. (d) Extracted vessel centerlines represented by polylines.

The process of reconstructing a superabundant 3D vessel model is shown in Figure 4.4. The polyline representing the vessel structure in the first fluoroscopic image is traversed, sampling at every 25 pixels. Since the projection geometry of the C-arm acquisition system is known (extracted from the system), epipolar lines of each point can be projected into the second image. All intersections between the epipolar line and the polyline in the second image are computed to generate multiple correspondences. Using the polylines in the second image instead of the centerline mask gives subpixel intersection accuracy. The points are triangulated to reconstruct the 3D vessel model. The structure of the resulting superabundant 3D vessel model is not noisy or unordered data. It has a vascular, tree like structure where some of the reconstructed branches are correct reconstructions and others are erroneous, as shown in Figure 4.4 (c). This reflects the structure of the vessels in the 2D images.

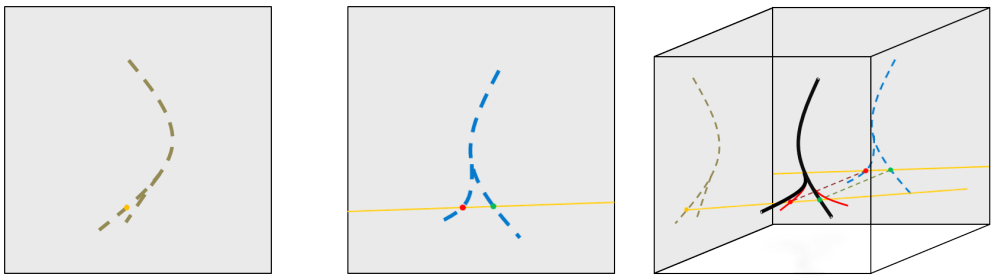
A ratio of correctly reconstructed points in the superabundant point cloud can be defined. The number of centerline (CL) points to reconstruct from the first view is known. It is assumed that every point in the first image has exactly one correct correspondence. The inlier ratio ϱ_{in} can be computed by dividing the number of points in the first image by the number of all epipolar correspondences in the second image:

$$\varrho_{in} = \frac{\text{\#points in 1st image}}{\text{\#correspondences in 2nd image}}. \tag{4.1}$$

It should be noted, even if the two sequences are gated successfully and the two frames are perfectly matched, the vasculature might be deformed differently in different cardiac cycles. This deformation can result in slight inaccuracies in the 3D reconstruction of vascular centerline points that can also affect registration accuracy.

4.2.4 Registering Adjacent Anatomy

Multimodal image alignment by registering adjacent anatomy is a novel solution to the problem of having few or no cross-modality landmarks. In CRT, the goal



(a) View one (b) View two (c) 3D reconstruction

Figure 4.4: Reconstruction of a vessel from two views. (a) View one with point to reconstruct (yellow). (b) View two with cast epipolar line (yellow), correct (green) and false correspondence (red). (c) Reconstruction of vessels from image pair, showing correctly (black) and incorrectly (red) reconstructed vessels.

is to register the LV epicardium to the coronary veins. This has two significant challenges: 1) The coronary vessels cover only a small fraction of the surface of the LV, i.e., it is a partial surface registration problem. Approaches to solve problems of this type are susceptible to falling into local minima. 2) The automatically reconstructed 3D venous model (Section 4.2.3.3) contains a large number of outliers. Individually, these problems can be difficult to solve. Together, they pose a substantial challenge. The proposed approach customizes and extends the GO-ICP algorithm [125, 126], making it robust to large number of outliers by dynamically setting the trimming factor.

4.2.4.1 Globally Optimal Registration of Partial Surfaces

The standard formulation of the registration problem is

$$E(\mathbf{R},\mathbf{t}) = \sum_{i=1}^{N} e_i^2(\mathbf{R},\mathbf{t}) = \sum_{i=1}^{N} ||\mathbf{R}\mathbf{x}_i + \mathbf{t} - \mathbf{y}_{j^*}||^2 , \tag{4.2}$$

where $\mathbf{x}$ represents the vessel point cloud, $\mathbf{y}$ the epicardial points, e_i is the error of point i depending on the rotation $\mathbf{R}$ and the translation $\mathbf{t}$, N represents the number of data points, and $\mathbf{y}_{j^*}$ the optimal correspondences.

The iterative closest point (ICP) algorithm [123] and its variants can be applied to solve this problem, however, ICP suffers from finding local minima

and not the global optimum. Alternatively, the branch and bound (BnB) algorithm finds the global optimum, but requires the whole search space to be processed making it computationally expensive [137, 138]. It has been shown that by combining these two methods, the global optimum can be found with reduced computational complexity [125, 126].

This approach is an encapsulation of two BnB algorithms and ICP to accelerate the optimization. The outer BnB algorithm operates on the rotation space SO(3), parameterized by the cube $[-\pi;\pi]^3$ and the inner one on the translation space $\mathbb{R}^3$, parametrized by $[-\xi;\xi]^3$, where π and ξ are the half side lengths of the initial cubes respectively. The lower bound for BnB is defined by

$$\underline{E} \doteq \sum_{i=1}^{N} e_i^2 = \sum_{i=1}^{N} \max(e_i(\mathbf{R}_{\mathbf{r}_0}, \mathbf{t}_0) - \gamma, 0)^2 \,, \tag{4.3}$$

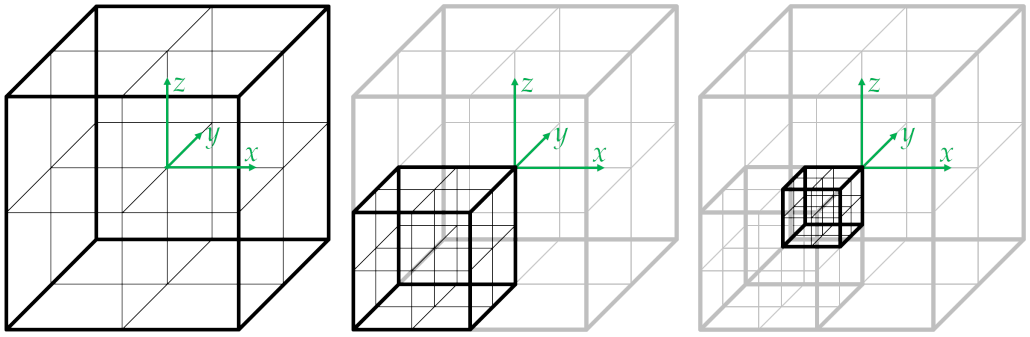
where $(\mathbf{r}_0, \mathbf{t}_0)$ represents the center of the current subspace $C_r \times C_t$ defined by the subcubes for rotation and translation respectively, $\gamma = \gamma_t + \gamma_r$ is the total uncertainty radius, that consists of the maximal distance in the current subcube from the center for the translation and the rotation respectively.

The algorithm subdivides the initial cube into octants (into eight subspaces), see Figure 4.5, and processes the subcubes in an order from the smallest lower bound to the highest. The upper bound is defined by

$$\overline{E} \doteq \sum_{i=1}^{N} \overline{e}_i^2 = \sum_{i=1}^{N} e_i^2(\mathbf{R}_{\mathbf{r}_0}, \mathbf{t}_0) \,, \tag{4.4}$$

that is equivalent to the point matching error at the center of the current subspace $(\mathbf{r}_0, \mathbf{t}_0)$. If the upper bound $\overline{E}$ is below the current best estimate E^*, ICP is called. If the current best error estimate E^* and the lower error bound $\underline{E}$ are within a preset threshold ϵ, $E^* - \underline{E} < \epsilon$, the optimal solution is found.

The LV and CS models are centered and scaled to be in the interval $[-1;1]$. Thus the range of translation can be limited to $[-1;1]$. The orientation is limited to be in the range of $[-35°;35°]$ which captures the clinically feasible range of rotations.



(a) Initial search space (b) Restricted space (c) Further restricted space

Figure 4.5: GO-ICP search space subdivision at the example of translation (same is valid for rotation). (a) Whole search space is subdivided into eight subcubes. (b) Most promising subcube is investigated and subdivided. (c) Subdivision is performed recursively for previous subcube.

4.2.4.2 Dynamic Outlier Trimming Factor

To cope with outliers resulting from the superabundant vessel reconstruction, the GO-ICP variant with trimming is used [126]. The trimmed GO-ICP algorithm amends the error bounds to use only a fraction of the vessel points for the current registration step and uses the trimmed ICP algorithm [139]. The fraction of points to exclude is called trim fraction and can be estimated by calculating the inlier ratio as described in Section 4.2.3.3 and inverting: $\varrho_{\text{trim}} = 1 - \varrho_{\text{in}}$.

The new lower bound is defined by the subset of points that are closest to the other point cloud in the current iteration as

$$\underline{E} \doteq \sum_{i=1}^{Q} \underline{e}_i^2 \,, \tag{4.5}$$

where Q is the number of points in the closest subset defined by ϱ_{trim}. The upper bound is defined by

$$\overline{E} \doteq \sum_{i=1}^{Q} \overline{e}_i^2 \,. \tag{4.6}$$

4.2.5 Imaging Parameters

The MR images are part of a standard cardiac protocol, non contrast-enhanced SAx and LAx steady state free precession (SSFP) cine images with a flip angle of

52° on a Siemens Aera 1.5 T MR scanner (Siemens Healthcare GmbH, Erlangen, Germany). The SAx stack has a resolution of 192 rows by 156 columns with a pixel spacing of 1.5625 mm and a slice thickness of 8 mm with a gap of 2 mm between slices. The LAx image is either a two, a three or a four chamber image. The LAx image was acquired with a resolution of 156 rows by 192 columns with the same pixel spacing as the SAx slices and a slice thickness of 6 mm. Both sequences were acquired over one cardiac cycle with 25 phases.

The X-ray images were acquired with a Siemens Artis interventional C-arm X-ray system (Siemens Healthcare GmbH, Forchheim, Germany). The images are sequences with 7.5 frames per second (FPS) and the imaging detector has a resolution of 1920×2048 pixels with a pixel spacing of 0.154 mm. The acquisitions are performed with different magnification factors and collimation. The sequences are acquired with a retrograde manual contrast agent injection with balloon occlusion to highlight the CS and its tributary veins.

4.3 Evaluation and Results

The presented method was evaluated on a phantom dataset, a clinical patient computed tomography angiography (CTA) dataset where a ground truth registration is available, and on nine clinical CRT cases. The experiments were performed for all datasets with known correspondence-based and superabundant vessel reconstructions.

4.3.1 Experiments

To evaluate the performance of the registration method, the same experiments were performed for the phantom and the clinical CTA dataset. The experiments were performed for both, the known corresponcence-based and the superabundant vessel reconstructions. The ground truth mesh was repeatedly perturbed by a rotation around a specific coordinate axis (x, y, or z). The rotation was in the range of -30 to 30° with 10° steps. No translation was applied, since the

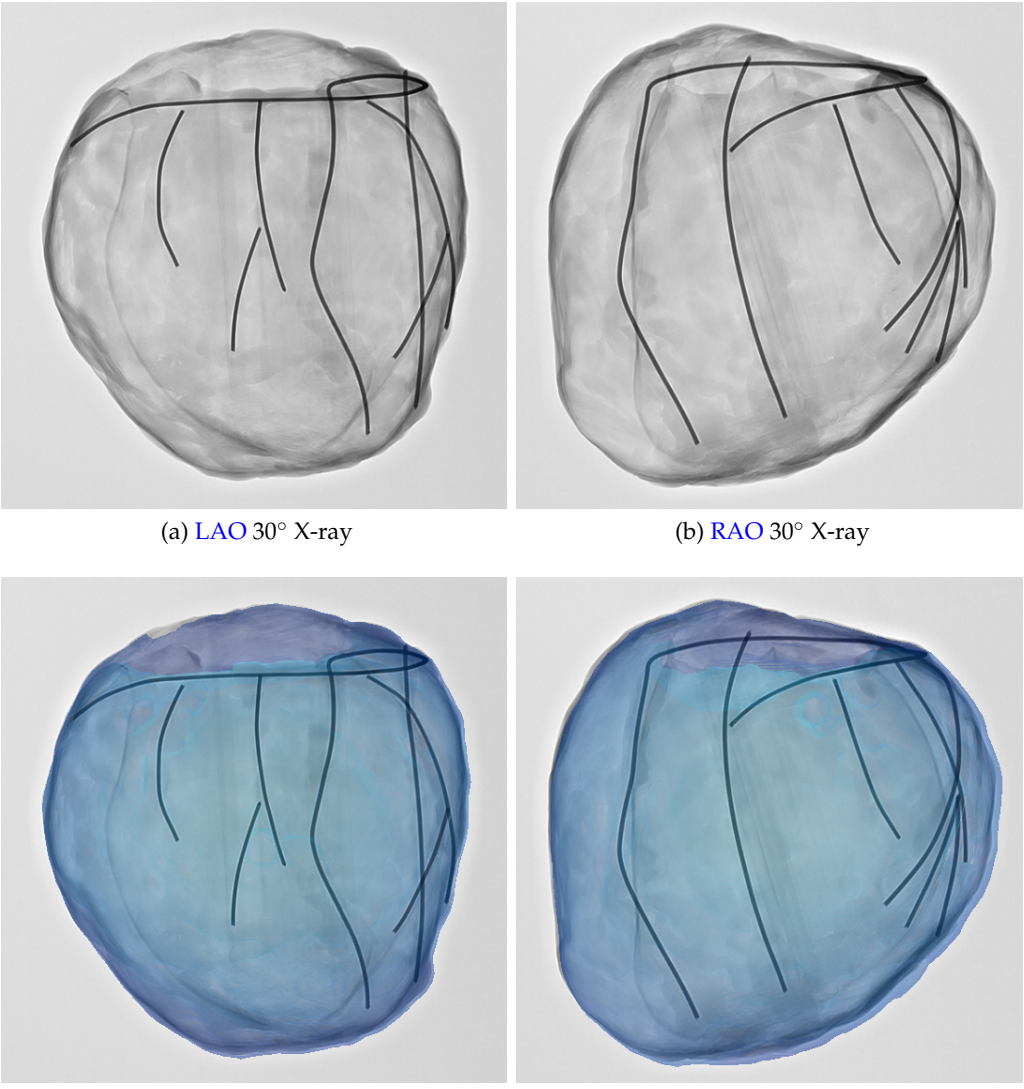
point clouds were centered as a first step of the registration as described in Section 4.2.4.1. The perturbed mesh was registered to the respective reconstructed vessel point cloud with the GO-ICP method. Registration was also performed with the ICP [123] and coherent point drift (CPD) [124] methods for comparison. The vertex-to-vertex mean absolute error (MAE) was measured.

4.3.2 Phantom Data

To be able to evaluate the performance of the method in a controlled environment with ideal data, a specially designed phantom was created. The epicardial shell of the LV of a CRT patient was segmented from a 3D whole heart MR dataset. The shell was 3D printed and metal wires were attached to model coronary veins. The model was imaged with the clinical C-arm X-ray system. A 3D dataset was acquired by a cone beam computed tomography (CBCT) acquisition. X-ray images at multiple angulations were recorded in the exact same pose. Finally, the acquired CBCT volume was manually segmented and a 3D surface model of the epicardium was extracted. The same pose guarantees that the X-ray images and the segmented epicardial model are in the same coordinate system, see Figure 4.6.

It is, however, to be noted that the registration is still not error-free. The residual error consists of segmentation inaccuracies and errors resulting from the 3D model generation from the segmentation. Small registration errors can occure, due to calibration inaccuracies of the CBCT reconstruction. To minimize the effect of calibration inaccuracies, the X-ray system was calibrated by a Siemens engineer. The slight errors of the ground truth registration are visible in Figure 4.6.

With the phantom dataset acquired, two experiments were performed. A registration experiment with known vessel correspondences, to prove that a registration is possible with the available data and a second experiment, with all epipolar correspondences reconstructed, thus with superabundant reconstruction.



(a) LAO 30° X-ray

(b) RAO 30° X-ray

(c) LAO 30° overlay

(d) RAO 30° overlay

Figure 4.6: Created phantom dataset. X-ray acquisitions of the phantom at (a) LAO 30° and (b) RAO 30°. Epicardial mesh, segmented from MRI, overlayed onto the two X-rays at (c) LAO 30° and (d) RAO 30°.

4.3.2.1 Registration with Known Vessel Correspondences

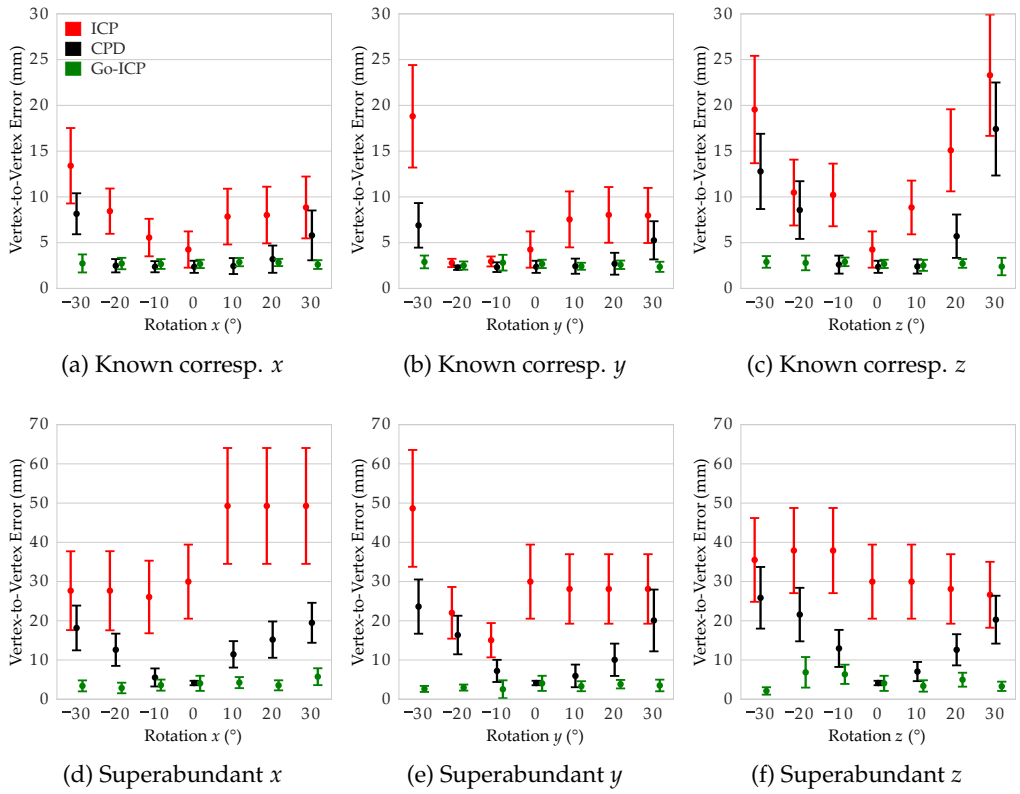
To evaluate the method under ideal circumstances, an experiment was conducted with a vessel reconstruction of known point correspondences. First, two images were selected having an angular difference of 60°, that is typical for cardiac (especially CRT) interventions, see Figure 4.6. Since the attached wires that represent the vessels are well contrasted, they can be easily separated from the background by thresholding. The thresholded binary images were skeletonized to extract the centerlines of the wires, as described in Section 4.2.3.2. The components of the centerlines were subdivided into separate images, such that only one centerline (or a section of a centerline) with known correspondence could remain in an image pair. This guarantees that for each point in the first image there is only one corresponding point in the second image, thus resulting in an ideal, outlier-free reconstruction.

The repeated rotation experiment (Section 4.3.1) with known correspondences showed, that the GO-ICP registration outperforms the other point cloud registration methods, see Figures 4.7 (a) to 4.7 (c). ICP might even fail if the rotation is small and the initial alignment is close to the optimal solution. ICP registration resulted in an average MAE of 9.54 ± 5.60 mm. CPD is more robust against rotation, often finds the optimal rotation for higher angles too. The CPD algorithm resulted in an average MAE of 4.90 ± 4.03 mm. The average MAE of GO-ICP was 2.68 ± 0.17 mm. This small error can be explained by the previously in Section 4.3.2 highlighted error factors.

4.3.2.2 Superabundant Point Cloud Registration

To simulate a scenario that could be applied to a clinical dataset, the registration experiment with the superabundant vessel reconstruction was performed, as described in Section 4.2.3.3.

Since the outliers disturb and prevent an accurate registration, GO-ICP with trimming was used. The trim fraction ϱ_{trim} was determined by the vessel point cloud reconstruction as described in Section 4.2.4.2. The estimated trim fraction



(a) Known corresp. x (b) Known corresp. y (c) Known corresp. z

(d) Superabundant x (e) Superabundant y (f) Superabundant z

Figure 4.7: Mean vertex-to-vertex errors and standard deviations of the registered phantom epicardial mesh with the ICP (red), CPD (black), and the GO-ICP (green) methods. (a-c) Errors of registration based on known vessel correspondences. (d-f) Errors of registration based on superabundant reconstruction. Note the difference in the error axis scales.

for the phantom dataset was 0.79. An experiment was performed where the trim fraction was varied and the resulting error of the automatically registered mesh was measured, see Figure 4.8. The experiment has shown that the error is minimal in the range of 0.80 to 0.85, thus in correlation with the estimated value.

The rotation experiment was performed and the registration results were compared to the ones of the registration with known vessel correspondences. The registration resulted in slightly worse mean errors for the superabundant reconstruction with an average MAE of 3.87 ± 1.22 mm. The GO-ICP method has clearly outperformed ICP (average MAE of 32.63 ± 9.52 mm) and CPD (average MAE of 13.26 ± 6.93 mm). ICP failed, even if the initial alignment was close to the global optimum. CPD performed well close to the optimum, but its performance decreased rapidly for higher rotations, see Figures 4.7 (d) to 4.7 (f).

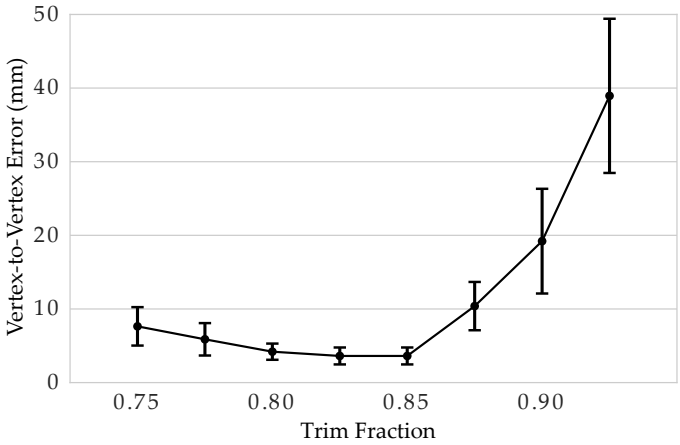


Figure 4.8: Error over trim fraction of the GO-ICP algorithm for the phantom dataset. For a rotation of 20° about the x axis.

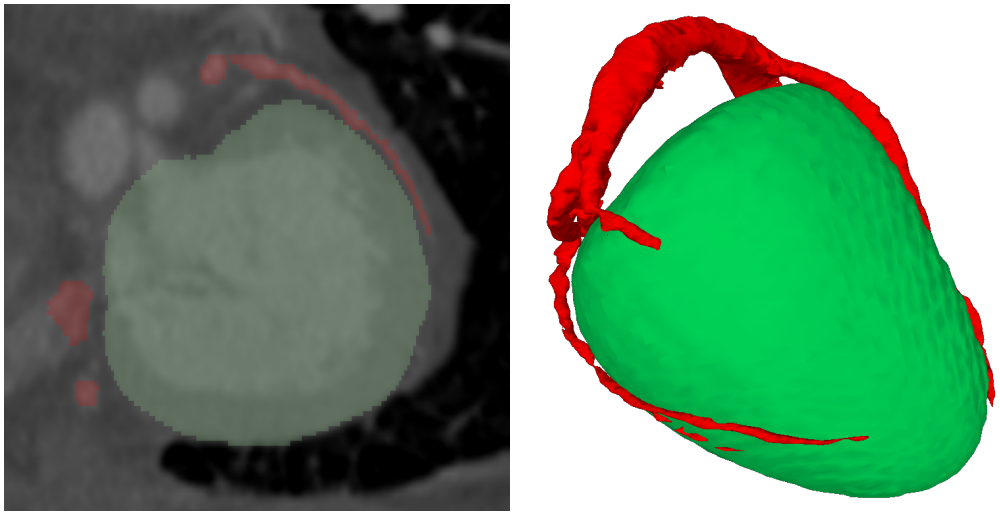
4.3.3 Clinical Data

After a successful evaluation on phantom data, the registration method was evaluated on clinical data. First, a cardiac CTA dataset was used, since a ground truth registration is initially available, thus it is possible to prove that the method is capable of registering ideal, real, clinical patient data. Additionally, the method was evaluated qualitatively on nine clinical CRT cases, on one of them quantitatively.

4.3.3.1 Clinical CTA Dataset

The evaluation of the method on clinical data is extremely challenging. Due to the lack of shared landmarks in the two modalities, an accurate ground truth registration is challenging to obtain and a target registration error is difficult to define. An accurate manual registration by a clinical expert is not feasible, since the registration with six degrees of freedom (DOF) cannot be easily performed with only two (non-orthogonal) X-ray projections, see Section 3.3.

To have a clinical dataset where a ground truth registration is available, a CTA dataset was segmented. Since both, the vessels and the LV, are extracted from the same image data, see Figure 4.9, both are initially registered, thus a ground truth registration is available and a vertex-to-vertex registration error can be calculated. The CTA segmentation has also revealed, what was described in



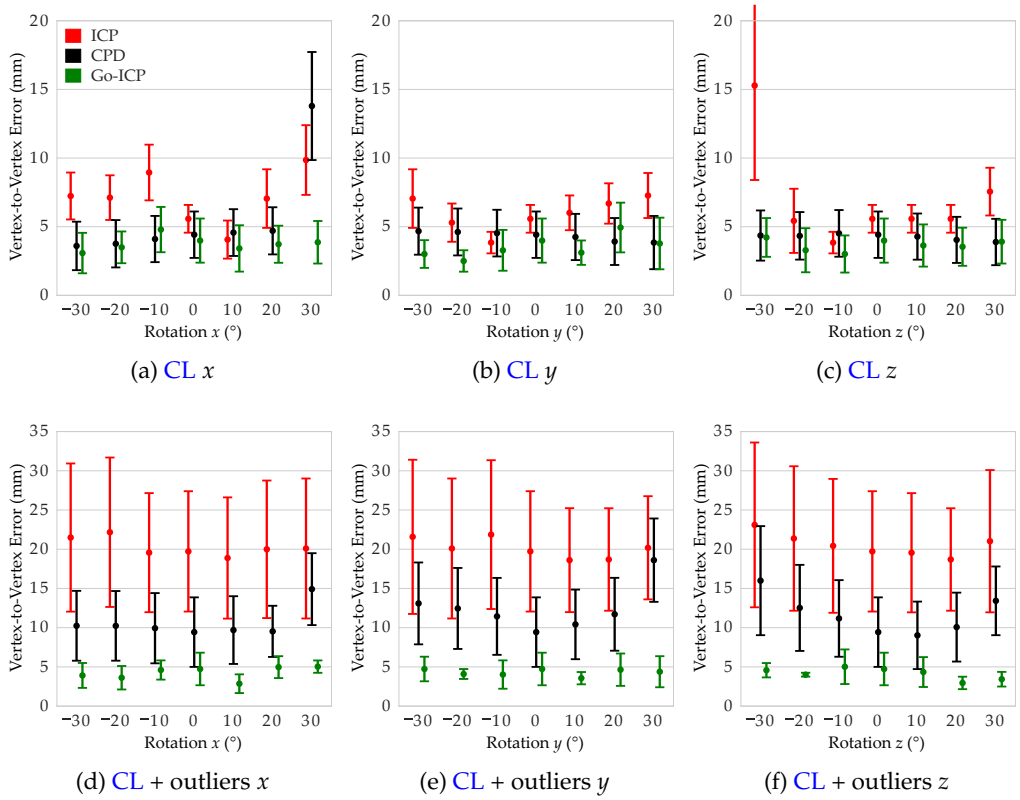
(a) CT + segmentation mask (b) 3D mesh models

Figure 4.9: Segmented CTA dataset. (a) LV epicardial mask (green) with coronary vein mask (red). (b) LV epicardial mesh (green) with the coronary vein mesh (red).

the literature [14, 140], that the whole CS does not always lie on the surface of the left ventricle, see Figure 4.9 (b). For this reason, the root of the CS was removed from the clinical vessel point clouds, because it might disturb the registration.

The same rotation experiment was performed for the clinical CTA dataset as for the phantom data, as described in Section 4.3.1. The GO-ICP method resulted in an average MAE of 3.65 ± 0.59 mm, thus, has outperformed ICP (6.69 ± 2.49 mm) and showed also marginally improved results compared to CPD (4.72 ± 2.10 mm), see Figures 4.10 (a) to 4.10 (c).

In a second experiment, to simulate a dataset with a superabundant vessel reconstruction, points of a uniform random distribution were added in the bounding box of the vessel centerline. The resulting superabundant point cloud had twice as many points as the initial point cloud, thus the trim fraction $\varrho_{\text{trim}} = 0.5$. Repeating the same experiment of rotations, GO-ICP could successfully register the epicardial mesh every time to the vessel point cloud, resulting in slightly worse results than in the case of the outlier-free, known correspondence-based point cloud, with an average MAE of 4.24 ± 0.65 mm. ICP failed in every case with an average error of 20.31 ± 1.24 mm. CPD could also not find the optimal



(a) CL x (b) CL y (c) CL z

(d) CL + outliers x (e) CL + outliers y (f) CL + outliers z

Figure 4.10: Mean vertex-to-vertex errors and standard deviations of the registered CTA epicardial mesh with the ICP (red), CPD (black), and the GO-ICP (green) methods. (a-c) Errors of registration for the reconstructed vessel CLs. (d-f) Errors of registration for the CLs with added outliers. Note the difference in the error axis scales.

alignment resulting in an average MAE of 11.56 ± 2.51 mm. For detailed results, see Figures 4.10 (d) to 4.10 (f).

4.3.3.2 Clinical CRT Datasets

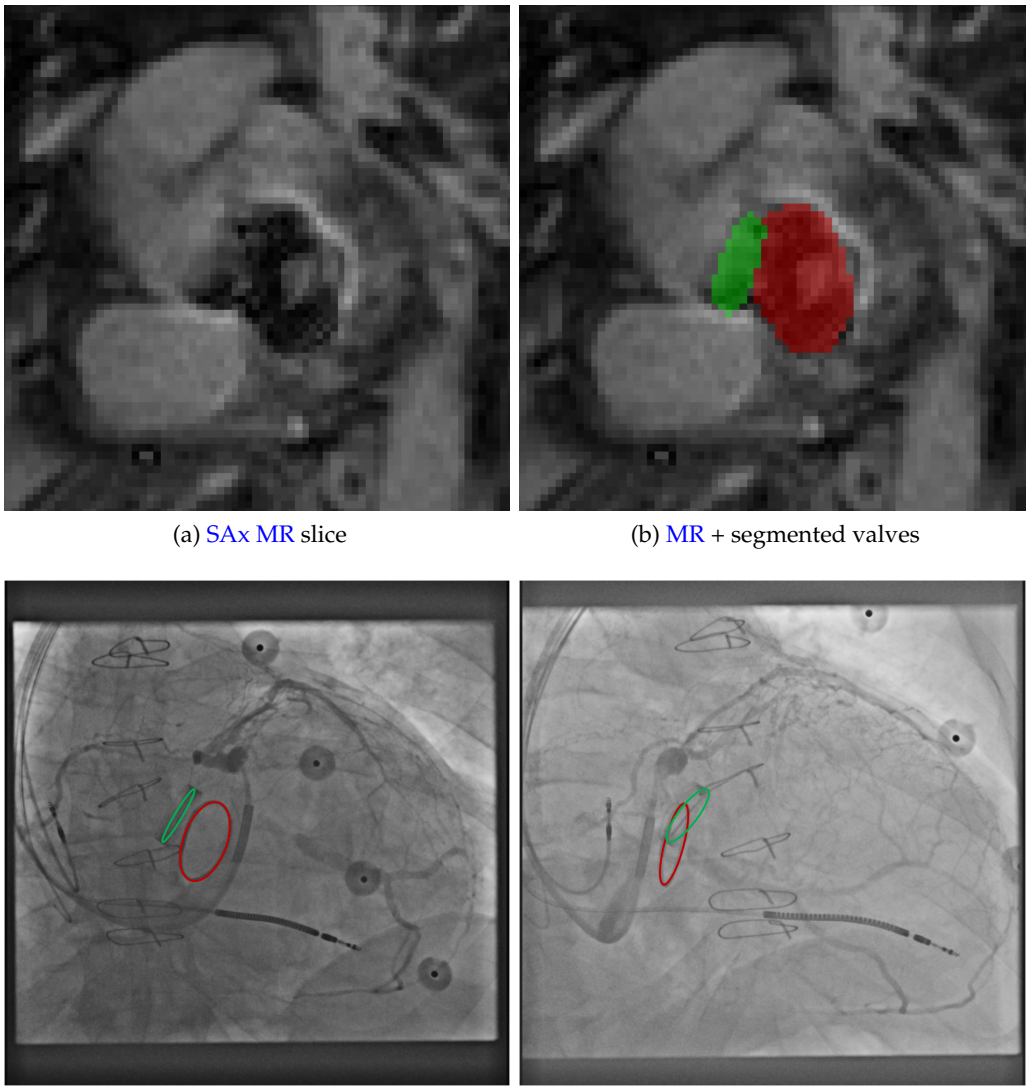
To be able to define an error measure for a real clinical CRT dataset, a novel evaluation approach, using previously implanted artificial valves, was developed. Among the available retrospective clinical datasets, one patient had previously implanted artificial aortic and mitral valves that are visible in both modalities, see Figure 4.11. The valves were segmented in the MRI dataset and their centerpoints were extracted in 3D. The two centerpoints were also reconstructed in 3D from the X-ray acquisitions. Thus, the extracted valve points can be used as multimodality landmarks. The valve points are ideal landmarks to have an estimation of the

registration error, since they are located directly at the base of the LV. Additionally, their arrangement defines the rotation of the LV around its long axis.

In a first experiment for the clinical CRT dataset, the registration with known vessel correspondences was performed. The results are visually appealing, see Figures 4.12 (a) and 4.12 (b), the overlay shows a good alignment with the heart shadow, the coronary veins appear to be on the ventricle and the location of the MRI derived valve centerpoints are close to the centerpoints is the X-ray images. The comparison of the valves in 3D resulted in valve errors of 4.29 mm and 9.13 mm for the aortic and the mitral valves respectively. Considering the resolution of the MR images used for the segmentation of the LV epicardial shell, 1.5625 mm $\times$ 1.5625 mm in-plane and 8 mm slice thickness with 2 mm gap between slices, the results are in the range of the slice spacing (slice thickness + gap = 10 mm).

To assess the accuracy of the results, clinical experts were consulted. The experts set the clinical requirement based on the 16 segment model, defined by the American Heart Association (AHA). The clinical team determined that the registration is sufficiently accurate, if the error does not significantly impact the target segment on the LV epicardium. The required accuracy was explicitly defined by the clinical team to be below half of the AHA segment size. The average segment size was found to be 43.4 mm (along short axis) $\times$ 32.7 mm (along long axis). Thus the minimum average half-segment size is 16.35 mm. The resulting valve registration errors satisfy this requirement, being below the average half-segment size.

In the second experiment, the epicardial mesh was registered to the superabundant point cloud. The trim fraction was calculated to be 0.44, thus $\varrho_{\text{trim}} = 0.44$ was used in the registration. The results are visually comparable with the results of the registration with known vessel correspondences, see Figure 4.12. This was also proven by the 3D valve errors. The aortic valve error was 2.94 mm and the mitral valve error was 3.86 mm, thus the errors differ only slightly from the known correspondence-based reconstruction's. The difference



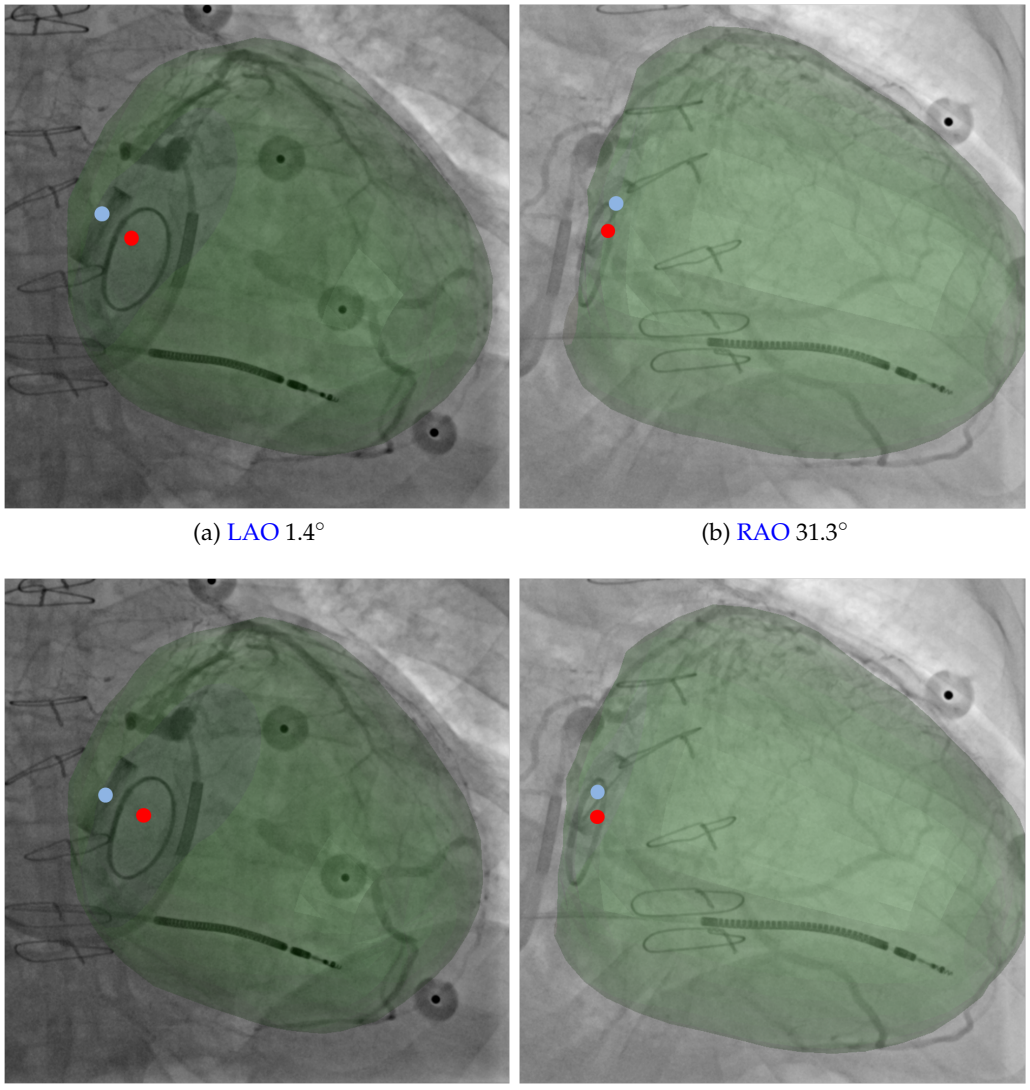
(a) SAx MR slice

(b) MR + segmented valves

(c) LAO 1.4° X-ray + valves

(d) RAO 31.3° X-ray + valves

Figure 4.11: Artificial valve extraction. (a) MRI slice showing artificial aortic and mitral valves. (b) MRI slice showing segmented aortic (green) and mitral (red) valves. (c) LAO 1.4° and (d) RAO 31.3° X-ray views with the aortic and mitral valves overlayed.



(a) LAO 1.4° (b) RAO 31.3°

(c) LAO 1.4° (d) RAO 31.3°

Figure 4.12: Registered epicadial mesh (green), aortic valve centroid (blue), and mitral valve centroid (red) overlayed. (a, b) X-ray views of registration with known vessel corrrespondences. (c, d) X-ray views of registration with superabundant vessel point cloud.

is in the range of the MR resolution.

The registration results are sufficiently accurate for the clinical use case of CRT (below the clinical requirement of 16.35 mm), however, the registration error could be further reduced. The accuracy of the system is affected by MR resolution and to a lesser extent MR segmentation and vessel reconstruction. The main limitation of accuracy is the very low out-of-plane resolution (10 mm) and low in-plane resolution (1.5625 mm) of the MR images compared with computed tomography (CT) which can have submillimetre voxels. Furthermore, this low resolution can introduce inaccuracies in the MR segmentation of the LV and manual annotation of the valves used for evaluation. Additional small sources of error can come from the segmentation and skeletonization of the vessels in X-ray and small inaccuracies in the phase gating between the two X-ray images which is limited by the frame rate (7.5 FPS). Inaccuracies in these steps can lead to 3D vessel reconstruction errors, however, this was not observed to be significant.

The registration with superabundant vessel reconstruction was evaluated on eight further CRT patients. Since no landmarks, such as artificial valves, were available, the registration was evaluated qualitatively by eight clinical experts, by visual inspection based scoring. The clinical experts assigned the overlay pairs of each patient to one of the four categories: 3. no correction is necessary for interventional guidance, 2. minor corrections are necessary, 1. major corrections are necessary, and 0. registration has failed, alignment is not useful for interventional guidance. This resulted in 72 ratings in total (8 experts $\times$ 9 patients). Scores of 3 or 2 were given in 95.8 % of cases, i.e., no, or only minor correction is necessary. Only 4.2 % of cases were rated with 1, requiring major correction. No failure was identified. The average rating of the registration was 2.31 ± 0.27. For a detailed summary of the results see Table 4.1.

The mean runtime for the nine CRT patients was 95.24 s on a workstation with an Intel Core i7 and 8GB RAM. The preprocessing time of the X-ray images and the runtime of the algorithm is acceptable for CRT procedures, since after the contrast agent injection was performed, the clinical team has to prepare for

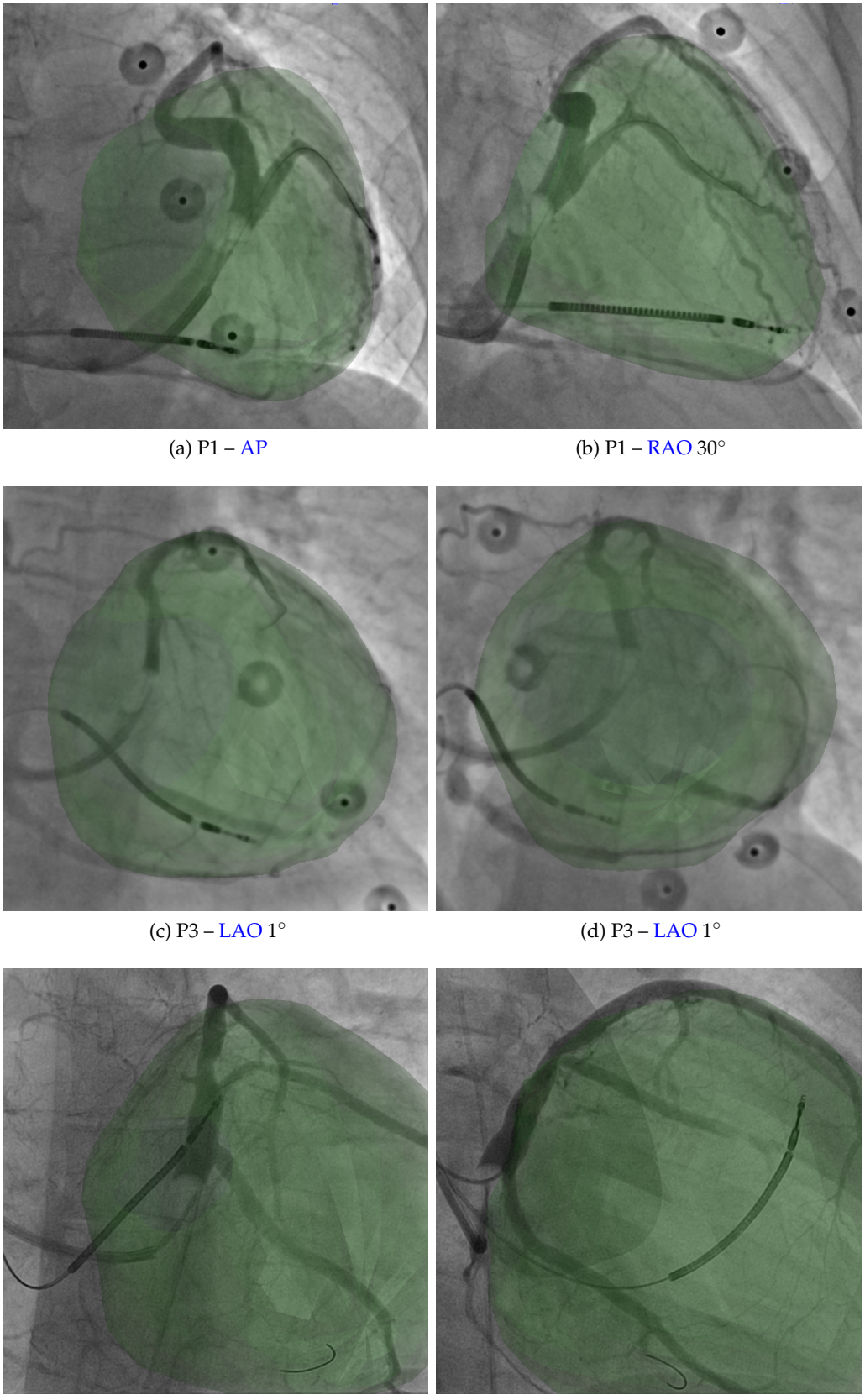
Table 4.1: Qualitative assessment of registration results by eight clinical experts for nine CRT patients. Scores: 3. no correction is necessary for interventional guidance, 2. minor corrections are necessary, 1. major corrections are necessary, and 0. registration has failed.

	Patients								
	P1	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**	**P9**
Expert 1	2	3	3	2	2	3	2	2	2
Expert 2	2	3	3	2	2	2	2	2	2
Expert 3	2	3	2	3	3	3	2	2	2
Expert 4	2	3	3	2	3	3	2	2	2
Expert 5	3	2	1	2	2	2	2	2	2
Expert 6	2	2	2	2	3	2	2	1	2
Expert 7	2	3	2	2	3	3	3	3	2
Expert 8	3	3	2	1	2	3	3	3	2
Mean	2.3	2.8	2.3	2.0	2.5	2.6	2.3	2.1	2.0
Std. Dev.	0.5	0.5	0.7	0.5	0.5	0.5	0.5	0.6	0.0

the LV lead implant. The occlusion balloon has to be removed and the catheter delivering the lead has to be prepared and inserted into the coronary sinus. It should be noted that the current implementation is single-threaded. A multi-threaded implementation could explore the search space in parallel, thus could greatly reduce runtime.

The overlay pairs for three patients are shown in Figure 4.13. The results are visually appealing, the models' projected epicardial borders corresponds well to the heart shadow, except for minor discrepancies. Additionally, the coronary veins appear mainly on the left ventricle in the images. Figures 4.13 (e) and 4.13 (f) show that the registration results in good visual alignment even for extensively cropped X-ray images.

It is to be noted that the registration appears to be robust against the non-simultaneous acquisition of the two images. Since the images are acquired with the same plane rotated, even after the frame matching was performed, as described in Section 4.2.3.1, the vasculature might have deformed between the acquisitions. These slight deformations can affect the reconstruction accuracy, but the results indicate the robustness of the method against them.



(a) P1 – AP (b) P1 – RAO 30°

(c) P3 – LAO 1° (d) P3 – LAO 1°

(e) P5 – LAO 18° (f) P5 – RAO 30°

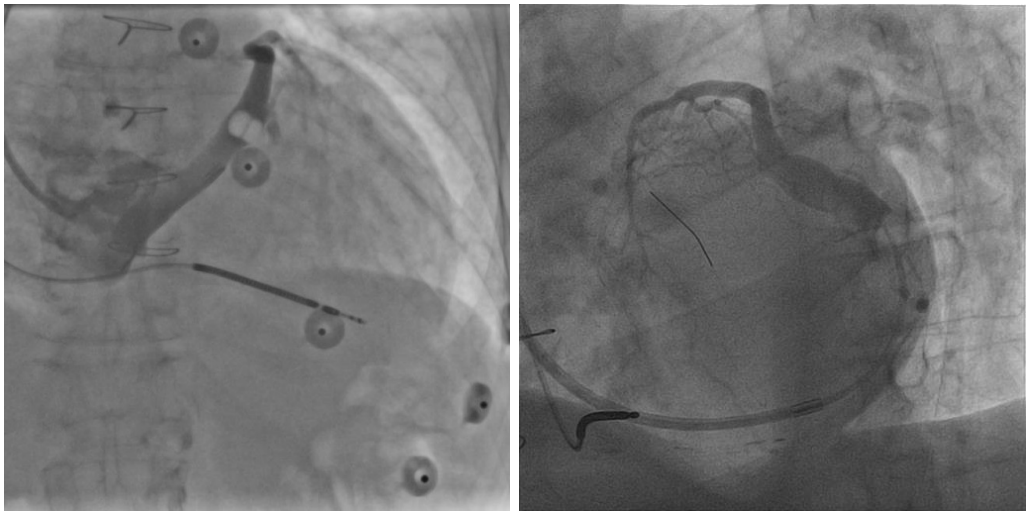
Figure 4.13: Overlay of the registered epicardial shell for three CRT patients.

4.4 Discussion and Conclusion

This chapter has presented a novel approach for registration in cardiac interventions and a method for evaluation by using implanted artificial valves. The idea is to use adjacent anatomical structures, the left ventricle segmented from cardiac MR images and a point cloud of the coronary veins that was reconstructed from two interventional X-ray images, without explicitly determining point correspondences, resulting in a superabundant point cloud. The reconstructed superabundant vessel point cloud has a high ratio of outliers, since all epipolar correspondences were reconstructed. For the registration of the epicardial model and the vessel point cloud, the GO-ICP algorithm with trimming for outlier rejection was used and the trim fraction was defined dynamically. Experiments on a specially designed phantom have shown that the method is capable of registering the two point clouds with an average MAE of 3.87 ± 1.22 mm that is superior to other point cloud registration methods, such as ICP and CPD.

Further experiments on a clinical CTA dataset have shown that the registration method is able to register clinical data. The average MAE between the ground truth and the registered epicardium was 4.24 ± 0.65 mm, slightly worse than for the phantom dataset.

In a final experiment the method was validated retrospectively on nine clinical CRT cases. The registration results were scored by eight clinical experts on a scale of 0 (worst) to 3 (best). The average score was 2.31 ± 0.27. One of the patients had previously implanted aortic and mitral valves that were segmented from the MR and X-ray images and reconstructed in 3D to define a registration error. After the registration, the valve centroids showed relatively low errors in 3D. The 3D aortic valve error was 2.94 mm and the mitral valve error was 3.86 mm. That is in the range of the MR slice spacing of 10 mm. The valve errors were below the clinical requirement defined by the half-segment size of a 16 segment AHA model, found to be 16.35 mm on average. The resulting valve errors suggest that the registration is capable of providing a registration that can be applied for image guidance of cardiac interventions, such as CRT delivery. The



(a) Insufficient balloon occlusion (b) Single major vein

Figure 4.14: Low quality venograms acquired in CRT interventions. (a) Insufficient balloon occlusion prevents segmentation and reconstruction of veins. (b) A single coronary vein does not constrain the point cloud registration sufficiently.

accuracy of the system could be further improved by using novel, high resolution MR protocols. Classical, 3D whole heart MRI acquisitions are not suitable, due to the long acquisition time and breath hold requirement. Free-breathing 3D MRI protocols are currently not widely available at clinical sites, in standard clinical workflows. However, as availability increases, the registration can potentially become an even more accurate element of the standard clinical workflow.

Although, it was shown that registering MRI to X-ray through exploiting the adjacent anatomical alignment of the LV and the coronary venous tree is feasible with sufficient accuracy for cardiac interventions, such as CRT, the method has its limitations. To reconstruct the coronary veins in 3D, two images with contrast injection in the veins (venograms) are required. The venograms have to be of sufficient quality that a 3D reconstruction is possible. This can be challenging to acquire, mainly due to the varying anatomy of the coronary veins throughout patients. If the CS is wide, the balloon occlusion for the retrograde injection may not be ideal and a backflow of contrast agent can be observed, see Figure 4.14 (a). This results in insufficient filling of the veins, preventing segmentation and 3D reconstruction, thus registration. The number of coronary veins can also

present variation, due to the high irregularity of coronary venous compared to coronary arterial anatomy. It can occur that a patient's left ventricle has a low number of coronary veins for 3D reconstruction, see Figure 4.14 (b). In such cases, registration can also not be performed successfully, because the reconstructed veins do not constrain the registration problem sufficiently.

To have a method that is not as strictly data-dependent as the adjacent anatomy-based registration, thus capable of registering data of patients even with inadequate coronary venous anatomy, an approach was developed that can register LV models to uncontrasted X-ray images. This novel method, relying on a modern machine learning approach, deep learning, will be described in more detail in the following chapter.

Chapter 5

Imitation Learning for Model-to-Image Registration

Chapter 4 introduced an adjacent anatomy-based approach for registering multi-modal data with no shared landmarks available. The approach is well aligned with the clinical workflow, but it is challenging to acquire ideal data during interventions, e.g., due to variations of individual patients' anatomy. In this chapter, a more generic, machine learning-based registration approach is presented. An imitation learning agent is trained purely on synthetic data to register mesh models to non-contrasted X-ray images. The imitation learning agent is evaluated qualitatively and quantitatively, on synthetic and clinical patient data.

5.1 Introduction

As described in Chapter 4, the adjacent anatomy-based registration with super-abundant vessel reconstruction can perform the registration of a preoperative mesh model to two interventional X-ray images, acquired with contrast agent injections. The method and the acquisition of the required data for the method do not interfere with the standard clinical workflow significantly. However, two X-ray images with contrast agent injections are required. The method is, however, limitedly applicable, if the contrasted X-ray acquisitions do not have sufficient quality, due to the anatomy or the contrast injection. Additionally, the inserted catheters and the inflated balloon can deform the anatomy of the coronary veins,

introducing errors into the registration process. The model-based property enables the usage of this method with any preoperative modality, if the required left ventricle (LV) model can be extracted.

Learning-based approaches that can be used for guiding procedures were also developed in recent years. A noteable approach registers a computer assisted design (CAD) model of an implant to X-ray images by a convolutional neural network (CNN) regression model [101]. However, the approach is difficult to generalize to anatomical data. It is only applicable to highly stiff objects of certain shapes, i.e., the implants. The rendering of the implant model is performed similarly to a previous approach [141].

Classical approaches often have low robustness and capture range. Uniform data and a good initial alignment is required. Novel machine learning-based approaches can overcome these challenges. An approach relying on artificial intelligence (AI) was shown to perform rigid 2D/2D and 3D/3D registration robustly on medical data [102]. In this approach, an artificial agent, modeled by an artificial neural network (ANN), is trained to learn a policy, i.e., an optimal strategy to take actions depending on the input images. Due to the high robustness of the approach, it is ideal to be applied in interventional guidance, where robustness may be more valuable than accuracy. The approach was extended to solve 3D/2D registration of the spine in computed tomography (CT) and fluoroscopy [55]. However, in this approach, the agent takes digitally reconstructed radiographs (DRRs) as input. DRRs can only be rendered for CT, thus the approach is not directly applicable to registration problems where the 3D modality is magnetic resonance (MR).

There are two significant challenges in AI-based cross-modality registration: 1) they require large sets of training data with ground truth (GT) registration and 2) they only work on the specific modalities and acquisition protocols they were trained on. The former is a significant problem for cardiac resynchronization therapy (CRT). Interventional fluoroscopy is, in general, not automatically stored. Patients may be imaged in modality-specific positions (e.g. arms up/arms down)

causing a non-rigid transformation and manually generating GT registration is time-consuming and inaccurate. The latter makes the registration systems vulnerable to changes in acquisition protocols and prevents general adoption of the same system for multiple clinical procedures.

In the pursuit of a general and robust cross-modality registration framework, it is proposed to exploit a byproduct of the preoperative diagnostic process – anatomical models, similarly to that in Chapter 4. In order to diagnose or characterize diseases, it is common to segment the anatomy of interest, i.e., the LV for CRT. The main advantage of using preoperative models is that the registration framework can be generalized, as it is independent of preoperative voxel intensities and acquisition parameters. The method can be trained on a single modality and applied to other modalities representing the same anatomy without retraining for specific cross-modality images.

In this chapter, a novel solution for multimodality registration for cardiac procedures is presented that has minimal interference with standard clinical routine. The approach is a combination of 3D model extraction from preoperative data and an artificial intelligence-based registration framework [55]. The system is capable of registering preoperative models to a single 2D X-ray image, without relying on voxel intensities or features from the preoperative modality. This means that the preoperative data can be of any modality (e.g., MR, CT, or ultrasound (US)), if relevant models can be extracted. The advantages compared to the adjacent anatomy-based approach presented in Chapter 4 are that: 1) the approach requires a single X-ray acquisition and 2) it does not rely on X-rays with contrast injections. As a further advantage, similarly to that in the adjacent anatomy-based approach, the model extraction is not an additional complication, since preoperative models are often created during preoperative planning and diagnostic reporting. Thus, the method can provide a robust registration for interventional guidance without major interference with the standard clinical workflow.

5.2 Methods

In the following sections, the methodology of imitation learning-based registration is described in detail. An overview of the concepts is given and the details of imitation learning for 3D/2D model-to-image registration are highlighted, i.e., the ANN architecture and its training details.

5.2.1 Overview

It is proposed to register models extracted from preoperative data, i.e., MR to intraoperative X-ray fluoroscopy, to guide cardiac interventions. An overview with a trained agent is illustrated in Figure 5.1. The 3D preoperative data is segmented prior to the intervention to extract a model of the anatomy of interest, i.e., the LV. During the intervention, an X-ray image (the fixed image) is acquired. A 2D projection image of the LV model is generated (the moving image) with the same imaging geometry as the X-ray image. The two images are processed by the imitation learning agent, modeled by an ANN, which predicts the reward for each possible action. The better the direction of an action, the higher the reward. The action with the maximum reward is chosen and is applied to the 3D model. The moving image is regenerated from the transformed model. These steps are iteratively repeated until convergence.

In the current setup, the registration is performed between a 3D model and a single fluoroscopy frame, not accounting for cardiac and respiratory motion in consecutive frames. The depth is assumed to be approximately correct after isocentering the volume and the X-ray image. The registration problem is restricted to the three degrees of freedom (DOF) defined by the imaging plane: x (horizontal) and y (vertical) translation and a rotation (around the axis of projection z).

5.2.2 Imitation Learning

The registration task can be formulated as a type of reinforcement learning problem [102], imitation learning. The agent's steps can be modeled as a Markov

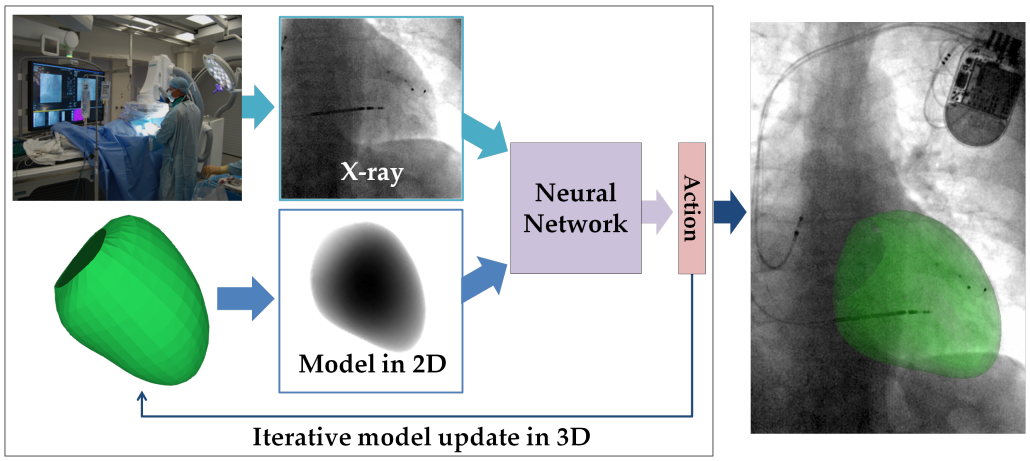


Figure 5.1: Overview of the model-to-image registration method with an artificial agent. An X-ray image is acquired during the intervention. The preoperatively extracted mesh model is projected into the imaging plane of the acquired X-ray image. A neural network (NN) predicts the optimal action to perform. The action is performed (transformation is applied to the model). This process is iteratively repeated until convergence.

decision process: $\{S, A, \tau, r, \gamma\}$, where S represents the possible states, A the possible actions, τ the probability of an action taken from a state at a certain time step, r the reward for an action, and γ the discount factor, defining the importance of long term rewards. The agent is in a single state (alignment) s_t at a certain time step t. Actions (steps) a_t in every direction along each DOF are rewarded, depending on the effectiveness (better or worse alignment) of action a_t. The goal is to learn a policy π, an optimal registration strategy that can predict the optimal action with the highest reward from the current state S_t:

$$a_t = \pi(S_t)\,, \tag{5.1}$$

thus to maximize the long-term reward:

$$\sum_{t=0}^{\infty} \gamma^t r_{a_t}\,, \tag{5.2}$$

where r_{a_t} is the reward for action a_t. The agent can be modeled by an ANN and, by training, a policy is learned by the network. The policy will imitate what the agent was being shown during training.

The agent is trained in a supervised manner: it is shown two images in the current state and the optimal rewards. The rewards are defined in a way that an action receives a higher reward, if it brings the moving image closer to the optimal alignment. The improvement, thus the reward r_{t+1}, is defined as the difference of distances between the old transformation T_t and GT transformation T_g, and the current transformation T_{t+1} and the GT transformation:

$$r_{t+1} = D(T_g, T_t) - D(T_g, T_{t+1})\,. \tag{5.3}$$

The distance between two transformations T_1 and T_2 is $D(T_1, T_2)$, the L2 norm of the parameters of the transformations, as described in [102].

5.2.3 Architecture

The agent is modeled by a pair of CNNs to encode the input images into features and another NN that decodes the features to determine the rewards, see Figure 5.2. The input layer of each CNN is defined to be 128×128, the input images are resampled to match this resolution. The CNNs consist of four convolutional layers. The layers have 16, 32, 64 and 128 filters, respectively. The filter size is 3×3 with a stride of 1×1 throughout. The convolutional layers are each followed by a rectified linear unit (ReLU) and a max-pooling layer of size 2×2. Batch normalization is applied after each layer. The CNNs result in feature vectors that represent the data. The feature vectors are concatenated and a NN with four fully connected layers decodes the feature vectors to predict the rewards. The first three layers have 512, 128 and 64 units. The number of units of the output layer depends on the number of DOFs. These are followed by ReLU layers and batch normalization.

5.2.4 Model-to-Image Registration

To train an agent for registration, perfectly aligned 3D models and 2D images are required. It is a significant challenge to acquire GT registrations for MR or US to X-ray data. Additionally, the number of available multimodal datasets is highly

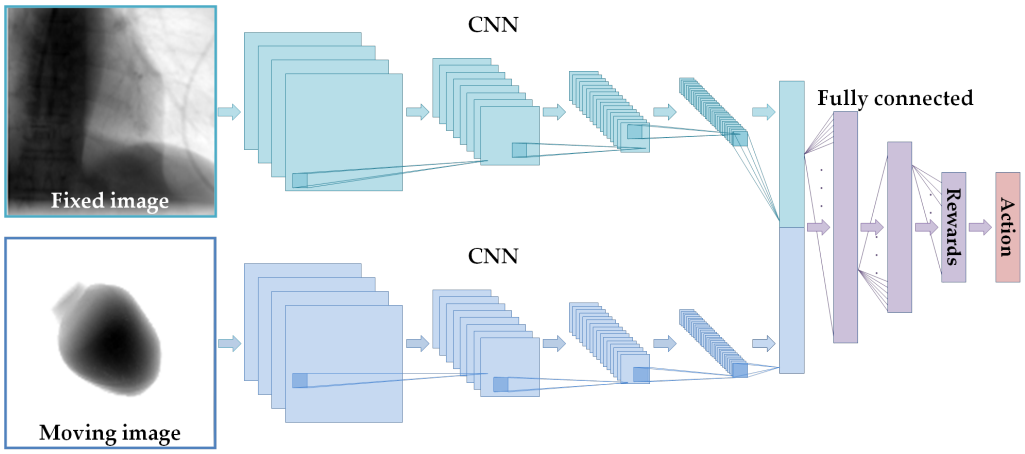


Figure 5.2: Architecture of the neural network that represents the imitation learning artificial agent. Each image, i.e., the fixed and the moving image, is processed by a CNN to extract features. The features are concatenated and fully connected layers predict a rewards for each possible action. The action corresponding to the maximal reward is chosen.

limited. Due to these reasons, exclusively CT images are used for training: the 3D models are extracted by segmentation and the 2D images are generated by projections.

The CT datasets are automatically segmented by a combination of object localization and a multi step non-rigid deformation estimation [121]. The segmentation results in a binary mask and a mesh model of the LV, see Figure 5.3.

5.2.4.1 3D/2D Registration Model

The problem of different dimensionalities is solved as in [55]: 2D images are shown to the agent. The fixed and moving images for every training sample are generated from the same CT dataset. The fixed image is a DRR [142] representing the intraoperative X-ray image. The DRRs are projection images of the CT volume, based on the X-ray attenuation model. The center of projection is defined to be the center point of the LV model. The fixed image is generated with a large field of view (FOV), i.e., 300 mm $\times$ 300 mm. The moving image, the projection of the LV model, is generated with a smaller FOV, i.e., 120 mm $\times$ 120 mm, having the LV centered. The model projection image corresponds to a subregion of the fixed image, the region of interest (ROI). Translation is performed by moving the LV

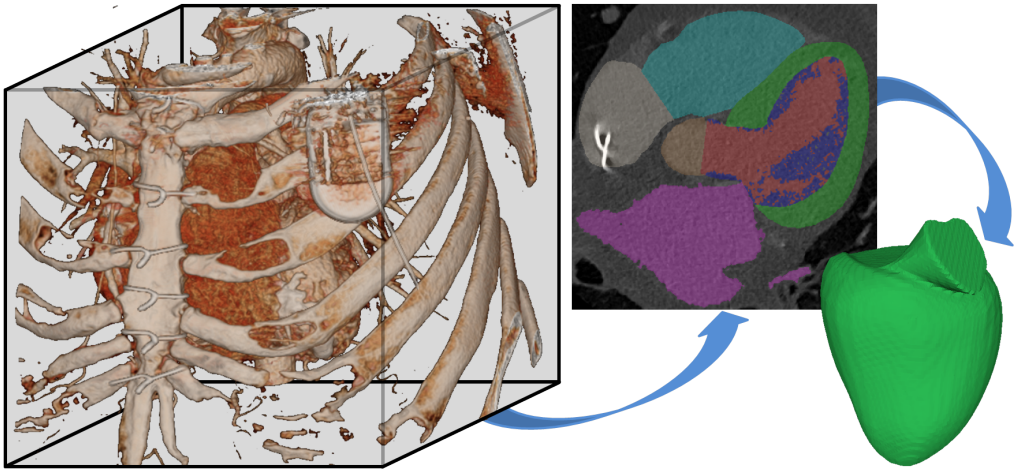


Figure 5.3: Model extraction from CT data. The CT volume (left) is automatically segmented. The 2D image slice (middle) shows the structures being segmented with colored labels. The 3D LV mesh model (right, green) is generated from the LV myocardium (green), the blood pool (red), and the trabeculations (dark blue).

model in 3D and regenerating the moving image, while keeping the LV model in the center of the FOV. This way, for consecutive translations, the projection image will correspond to different subregions of the fixed image. The ANN, modeling the agent, is shown an image pair: the moving image and the corresponding ROI extracted from the fixed image.

5.2.4.2 Training Setup

The LIDC-IDRI public dataset [143] and previously acquired data were used for training. The datasets cosist of 802 contrasted CT volumes in total. The data was split to 702 training and 100 test datasets. To generate a sufficient number of training pairs, the 702 training datasets were augmented. This was mainly performed by perturbing the perfectly aligned, generated image pairs. Transformations with the three DOFs (two translations and one rotation), defined by the imaging plane of the fixed image, were applied to the 3D mask.

The maximal perturbation of the translation components of the GT transformation was 35 mm and the maximal rotation component was 15°. These values correspond to misalignment observed in the clinical setting, after the isocenters of an MR volume and a fluoroscopy image are co-registered. Furthermore, a

random, maximally 10 mm offset to the center of projection was introduced, since the heart is not perfectly centered in fluoroscopy acquisitions. Additionally, the primary positioner angulation, i.e., right anterior oblique (RAO)/left anterior oblique (LAO), was varied between -15° and +15° around the anterior-posterior (AP) 0° projection, and the secondary angle, i.e., caudal (CAU)/cranial (CRA), between -5° and +5°. By generating 1000 perturbations for each of the 702 training datasets, 702 000 perturbations were created.

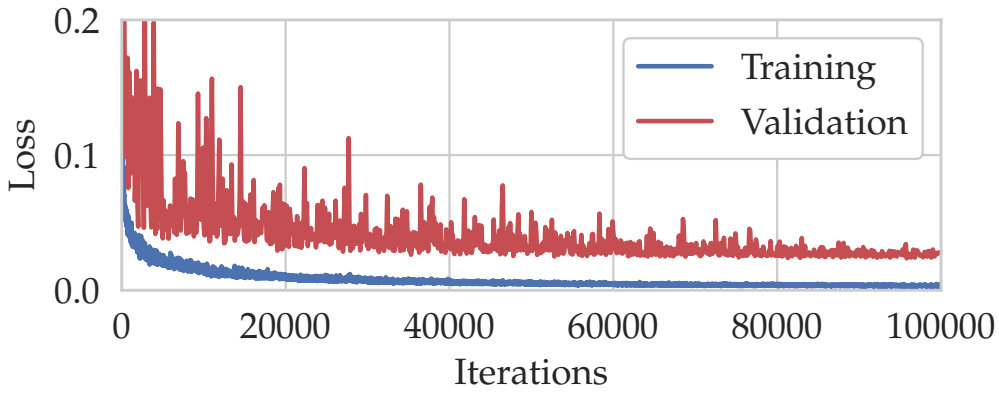
The same process was repeated for the other two common angulation ranges used in CRT: RAO 30° ± 15° and LAO 30° ± 15°. Thus, three individual networks were trained in total, one for each angulation range. This results in more specialized networks of higher performance for the respective angulation ranges.

The network described in Section 5.2.3 was trained with a minibatch size of 80. The solver used was RMSProp [144] with a momentum of 0.9 and the learning rate was 0.01 with a decay ratio of 0.8 after every 10 000 iterations. All weights were initialized from a uniform distribution generated based on the approach of He et al. [145] and the biases with the constant value of 0.1. The number of iterations performed was 100 000, corresponding to 11.39 epochs to ensure convergence. Training took approximately 20 hours on an NVIDIA GeForce GTX Titan X Pascal GPU.

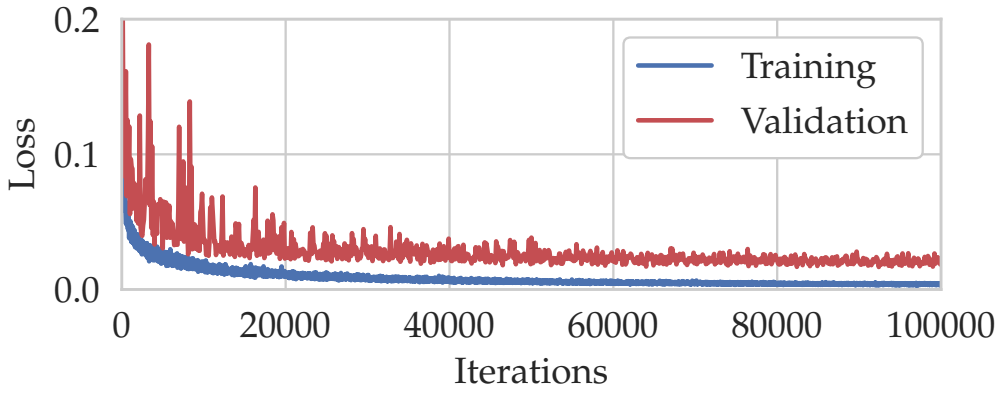
The training curves are shown in Figure 5.4. It is to be noted that the validation error (red) is more unstable than the training error (blue), but it converges for all three angulations to approximately 0.025. Furthermore, the training and validation curves do not reach the same loss level. The gap indicates that the model is slightly overfitting to the training data.

5.3 Evaluation and Results

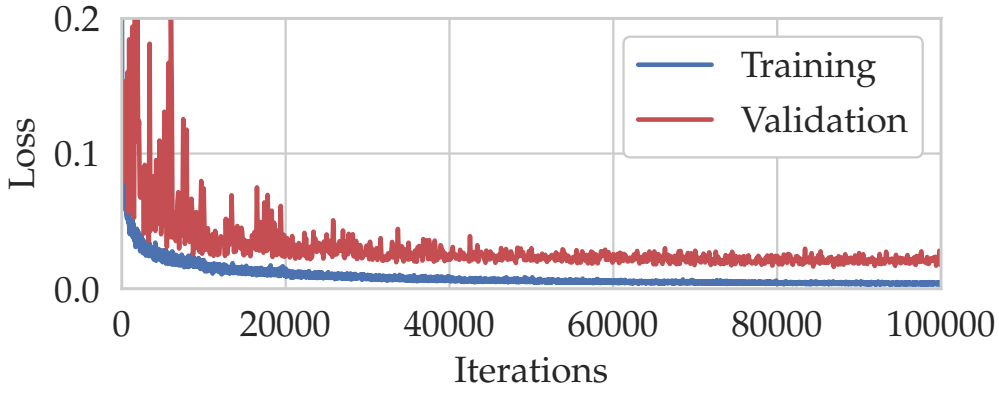
The evaluation of the imitation learning-based registration agent was performed on synthetic data for accuracy and on clinically acquired CRT data for robustness against geometric perturbations.



(a) AP range



(b) RAO range



(c) LAO range

Figure 5.4: Training (blue) and validation (red) curves for the (a) AP, (b) RAO, and (c) LAO angulation ranges for 100 000 iterations.

5.3.1 Synthetic Data

As described in Section 5.2.4.2, the data was split into 702 training and 100 test datasets. Each test dataset was perturbed ten times for each of the three angulation ranges (AP, RAO, LAO), resulting in 3000 test cases in total (1000 respectively). First, an experiment was performed to demonstrate that it is beneficial to train three individual, more specialized networks for the angulation ranges. Then the registration performance of the three specialized networks was evaluated qualitatively and quantitatively.

5.3.1.1 Clinical CRT Ranges vs. Global Angulation Range

It can be beneficial to train multiple networks for performing a task on data that can be subdivided based on a meta-parameter. In the case of imitation learning-based cardiac registration, this can be the angulation that the X-ray images are acquired at. As described in Section 5.2.4.2, the three standard CRT angulations were chosen: AP, RAO 30°, and LAO 30°. For all three angulations, a range of $\pm$ 15° was defined. Additionally, a single network was trained on the global angulation range of AP $\pm$ 45°, encompassing all possible CRT angulations. The three networks trained on the clinical ranges (CRs) and the network trained on the global range (GR) were comparatively evaluated on the synthetic test data of the three ranges for registration accuracy, see Table 5.1.

The results show that the CR networks generally outperform the network trained on the GR. In the LAO range the differences are not very distinct. The results are comparable in most measures. The standard deviation and the maximum error are lower with the network trained on the GR.

5.3.1.2 Qualitative Evaluation

To evaluate the method qualitatively, the projections of the LV model were compared with the corresponding fluoroscopy images after registration, see Figure 5.5. The only visual cue inherently found in the fluoroscopy image is the shadow of the left ventricle. Additionally, a cross shaped landmark is defined at the center

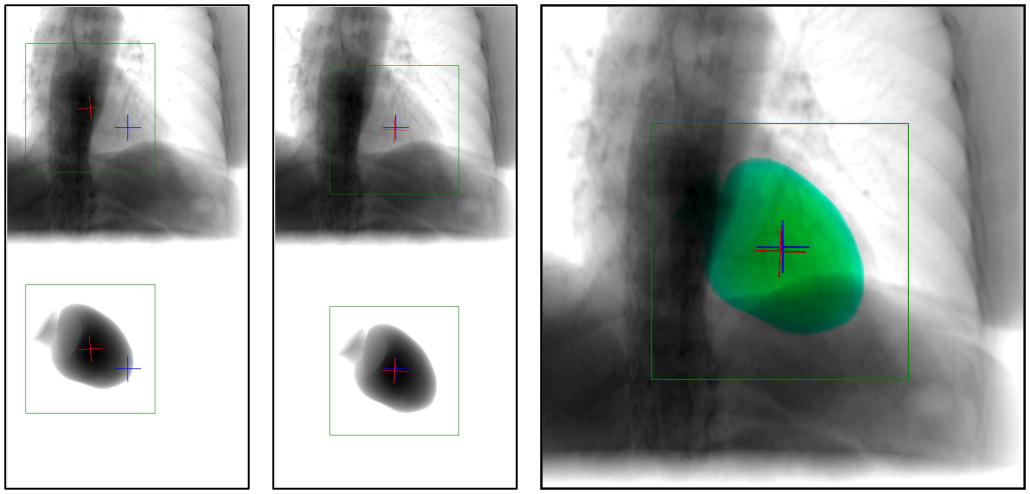
Table 5.1: Target registration error (TRE) of the cross landmark on synthetic data initially (Start) and after registration in mm for the global range (GR) and the three clinical ranges (CRs): AP, RAO, and LAO.

			Percentiles					
	Mean	**StD.**	50 %	60 %	70 %	80 %	90 %	100 %
AP – Start	22.80	10.50	21.42	25.22	30.03	33.50	36.96	47.88
AP – GR	3.12	2.52	2.43	2.89	3.40	4.34	6.07	19.61
AP – CR	2.92	2.22	2.34	2.80	3.45	4.23	5.76	16.11
RAO – Start	22.71	10.70	21.36	25.74	30.17	33.67	36.86	48.45
RAO – GR	4.06	3.40	3.09	3.75	4.68	6.20	8.02	25.25
RAO – CR	3.50	2.79	2.74	3.35	4.02	5.28	7.09	20.09
LAO – Start	22.49	10.43	21.51	25.88	29.29	32.88	36.43	47.37
LAO – GR	2.72	1.86	2.32	2.70	3.19	3.92	5.03	13.36
LAO – CR	2.67	1.99	2.16	2.58	3.00	3.63	4.98	17.79

of the LV, computed from the model of the LV. The cross extends 10 mm in every DOF from the center point. In successful registrations, the shadow of the left ventricle in fluoroscopy matches the border of the projected LV model and the landmarks are located at the same location, having the same orientation in both images, see Figure 5.5 (c).

5.3.1.3 Quantitative Evaluation

The TRE was measured by computing the L2 norm of the points of the cross landmark at the center, between the GT DRR cross (blue) and the registered LV model cross (red), as described in Section 5.3.1.2, see Figure 5.5. The TRE was computed in 2D, because the registration is performed in-plane. The depth is not adjusted, thus the 3D error would not provide more information. The quantitative evaluation was performed on the three angulation ranges. In more detail on the AP range and comparatively on the RAO and LAO ranges.



(a) Initial alignment (b) Final alignment (c) LV overlay at final alignment

Figure 5.5: Relation of fixed and moving image (a) before, (b) after registration, and (c) the overlay of the registered mask (green). Showing the ROI (green box), the fixed (blue cross), and the moving image landmark (red cross).

AP range

On the AP angulation range, the imitation learning agent was evaluated against manual and gradient-based automatic registration, see Table 5.2. The gradient-based metrics were gradient correlation (GC), gradient information (GI), gradient orientation (GO) [146], and their versions utilizing only the positive gradients (GC+, GI+, GO+), corresponding to the visible heart shadow in the images. This can eliminate strong gradients resulting from the liver dome, the diaphragm, or from the vertebrae of the spine.

The imitation learning agent's results were significantly better than those of the other approaches. The starting TRE of 22.8 ± 10.5 mm was improved to 2.92 ± 2.22 mm, the median TRE was reduced from 21.42 mm to 2.34 mm, and the angular error from $7.17 \pm 4.64°$ to $2.34 \pm 1.87°$. The best gradient-based method, GI+, has resulted in a TRE of 6.79 ± 4.75 mm with a median of 5.63 mm and an angular error of $7.28 \pm 4.71°$, showing slightly lower accuracy than manual registration (mean: 6.48 ± 5.60 mm, median: 4.93 mm, angle: $6.21 \pm 5.17°$).

The main reason for failures in gradient-based methods was that the highest metrics score is at the liver dome or the spine, providing the strongest gradients

Table 5.2: TRE of the cross landmark initially (Start) and after registration in mm.

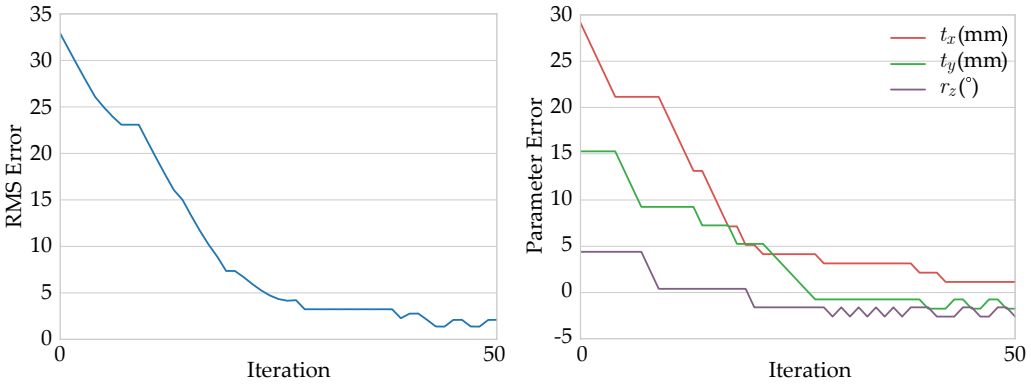
			Percentiles					
	Mean	**StD.**	50 %	60 %	70 %	80 %	90 %	100 %
Start	22.80	10.50	21.42	25.22	30.03	33.50	36.96	47.88
GO	9.65	6.23	8.39	9.80	11.50	14.08	17.78	46.83
GO+	10.49	5.97	9.42	10.87	12.49	15.00	18.54	38.02
GC	9.15	6.74	7.74	9.32	11.03	13.68	18.12	44.09
GC+	7.80	6.30	5.91	7.51	9.30	11.55	16.43	48.37
GI	8.44	6.61	6.47	7.58	8.97	11.68	16.37	48.55
GI+	6.79	4.75	5.63	6.50	7.48	8.84	11.77	46.14
Manual*	6.48	5.60	4.93	5.97	7.49	8.70	11.37	40.82
Agent	2.92	2.22	2.34	2.80	3.45	4.23	5.76	16.11

*Note: manual registration for a single, randomly chosen perturbation in each case.

in the DRRs. The amended methods (GO+, GC+, GI+) counteract this by using only positive gradients. These mainly correspond to the heart shadow that is usually visible in X-rays (and the generated DRRs), and the overlap with other structures, i.e. the liver, is minimal. This has improved the results in the metrics GC and GI.

A further complication is that in many cases the heart shadow is faint or blurry. This is the main reason for lower accuracy than with the agent-based approach. The agent can leverage multiple, non-hand-crafted features. It does not have to rely only on the gradient information, thus can register reasonably well, even in low quality data. It has improved the misalignment in every case. The results are promising, showing an improvement compared to current techniques. This suggests that the technique could be employed in cardiac interventions, such as CRT.

The evolution of the TRE and individual parameters of the agent are visualized in Figure 5.6 for the case shown in Figure 5.5. The TRE decreases monotonously until convergence. The figures show that a well trained agent's



(a) RMS TRE (b) In-plane parameter error

Figure 5.6: Evolution of the (a) root mean square (RMS) TRE and (b) individual parameters in the case shown in Figure 5.5. Coordinate axes are defined by the imaging plane: t_x and t_y being the horizontal and vertical translational misalignments respectively. The in-plane rotation is represented by r_z.

actions converge monotonously to the optimal alignment. A registration is performed within 3 s.

RAO and LAO Ranges

The imitation learning agent was also evaluated on the other two common CRT angulation ranges: RAO and LAO, see Table 5.3. The agents trained on these two ranges were evaluated against manual registration, the gradient- and classical optimization-based approach with the highes performance on the AP range (GI+), and to the results of the AP range. It is to be noted that different perturbations were applied to the three ranges, thus the ranges are only limitedly comparable with each other.

Agents trained on both ranges, i.e., RAO and LAO have superior performance compared to manual or gradient-based registration. It is to be noted, that manual registration results on LAO are comparable to those on AP, whereas results of manual registration on RAO show lower performance than on AP. This can be explained by the shadow of the heart partially originating from other chambers of the heart, i.e., the right ventricle (RV), right atrium (RA), and the left atrium (LA), see Figure 5.7. Due to this, the ground truth registration does not

Table 5.3: TRE of the cross landmark on synthetic data initially (Start) and after registration in mm for the three angulation ranges: AP, RAO, and LAO.

			Percentiles					
	Mean	**StD.**	50 %	60 %	70 %	80 %	90 %	100 %
AP – Start	22.80	10.50	21.42	25.22	30.03	33.50	36.96	47.88
AP – GI+	6.79	4.75	5.63	6.50	7.48	8.84	11.77	46.14
AP – Man.*	6.48	5.60	4.93	5.97	7.49	8.70	11.37	40.82
AP – Agent	2.92	2.22	2.34	2.80	3.45	4.23	5.76	16.11
RAO – Start	22.71	10.70	21.36	25.74	30.17	33.67	36.86	48.45
RAO – GI+	8.58	5.79	7.41	8.65	9.95	11.63	15.29	45.65
RAO – Man.*	9.12	8.40	6.55	7.95	10.70	13.21	16.51	53.47
RAO – Agent	3.50	2.79	2.74	3.35	4.02	5.28	7.09	20.09
LAO – Start	22.49	10.43	21.51	25.88	29.29	32.88	36.43	47.37
LAO – GI+	5.62	3.84	4.52	5.21	6.22	7.51	10.03	36.84
LAO – Man.*	6.64	7.04	5.08	6.07	7.46	8.87	11.10	57.40
LAO – Agent	2.67	1.99	2.16	2.58	3.00	3.63	4.98	17.79

*Note: manual registration for a single, randomly chosen perturbation in each case.

always show good alignment with the heart shadow, the main landmark used by a human in the registration.

A similar tendency can be seen for the gradient-based (GI+) method. The performance on the LAO range is slightly better than on the AP range and it is the worst on the RAO range. This can be explained by the previously described fact that the heart shadow is comprised of not only the LV. The heart shadow's strong gradients guide the optimization minimally away from the optimal alignment. It is to be noted that the performance of the gradient-based approach was higher than that of manual registration.

The RAO and LAO agents' performance was compared against the agent trained on the AP angulation range. The performance was slightly higher on the LAO and slighly lower on the RAO than on the AP range. This is in correlation with the results of the manual registration. The most plausible explanation is that the other chambers contribute to the heart shadow, resulting in weaker, more

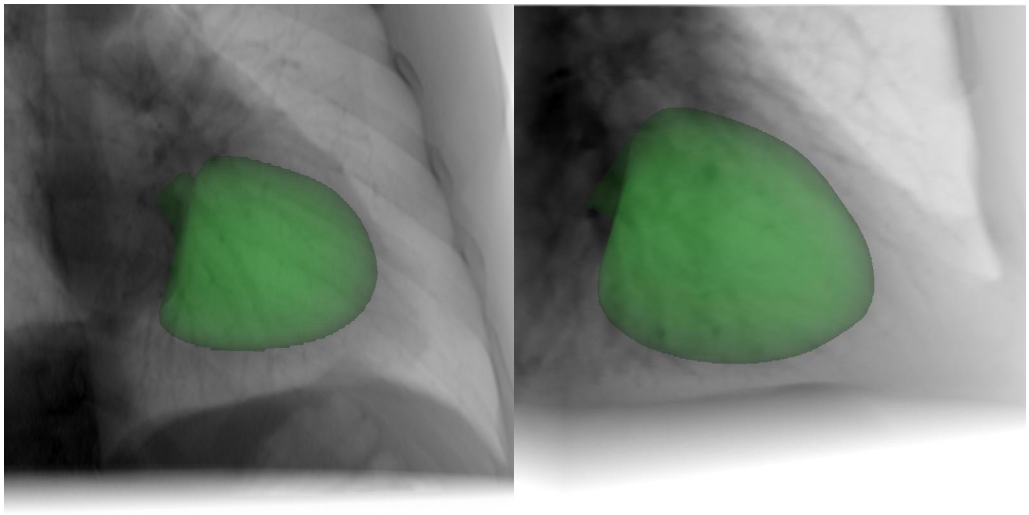


Figure 5.7: GT alignment of the LV mesh model in two exemplary RAO DRR projections. The silhouette of the LV model projection in 2D does not necessarily align with the heart shadow in the DRR projection. Other chambers, the RV and the atria, can contribute to the shadow as well.

limited landmarks.

5.3.2 Clinical CRT Data

Further evaluation was performed on 19 clinical CRT datasets to evaluate registration performance in a realistic scenario. The patient datasets consist of a combination of a segmented mesh model from an MR acquisition and a combination of X-ray images of AP, RAO, and LAO angulations. Each of the 19 datasets has one AP X-ray. RAO and LAO images were not acquired for all patients in that the whole LV was visible: 12 patients had images in the RAO and 10 patients an

Table 5.4: The available cases for the three angulation ranges AP, RAO, and LAO of 19 clinically acquired CRT patient datasets. Individual cases are missing from the RAO or LAO ranges, because images of those ranges were not acquired in the respective procedures.

	MR Guided CRT Case Number																		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**
AP	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
RAO	–	–	1	2	3	4	–	5	6	7	–	8	9	10	11	–	–	–	12
LAO	–	–	–	1	–	2	–	3	4	5	–	6	7	8	–	–	9	10	–

image in the LAO range. In the following, the cases are numbered based on the available images in each angulation range, i.e., case 5 in AP is not from the same patient as case 5 in RAO. For an overview of available images in the individual angulation ranges, see Table 5.4. The secondary angulation (CRA/CAU) was in the range of $0° \pm 5°$.

The end-diastolic X-ray frame is manually selected. Corresponding end-diastolic models were extracted from the MR images by a combination of a machine learning-based landmark detection and a minimum path algorithm based on histogram analysis [129].

An accurate ground truth registration is not available, thus accuracy was evaluated qualitatively. Since in cardiac interventions, such as CRT, robustness of registration has priority over accuracy, the method was evaluated for robustness. After a rough, initial manual alignment, the models in the 19 cases were perturbed multiple times by in-plane, three DOF transformations for each of the angulation ranges, similarly to how it was performed on the training data. The perturbed models were registered to the corresponding X-ray images by the imitation learning agent. If the registration provides similar results for different perturbations, the method is robust against rigid geometric perturbations. High robustness ensures that the registration can be successful from different starting positions, i.e., after isocentering the model and the X-ray. Robustness was evaluated qualitatively and quantitatively.

5.3.2.1 Qualitative Evaluation

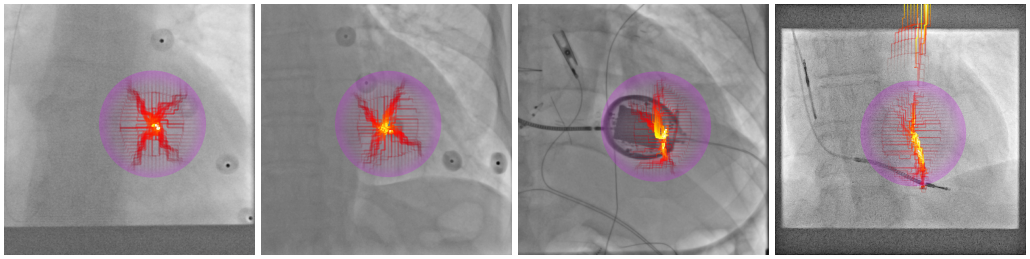
Qualitative evaluation was performed through visualizing the agent's trajectories from the starting to the final positions. The LV model was perturbed 100 times from an initial rough manual alignment. Misalignments of the center of the LV were on an equally sampled circle with a radius of 30 mm. The perturbations were generated for the three angulation ranges: AP, RAO, and LAO. The perturbed models were re-registered to the X-ray images and the path of the center of the LV was recorded.

AP range

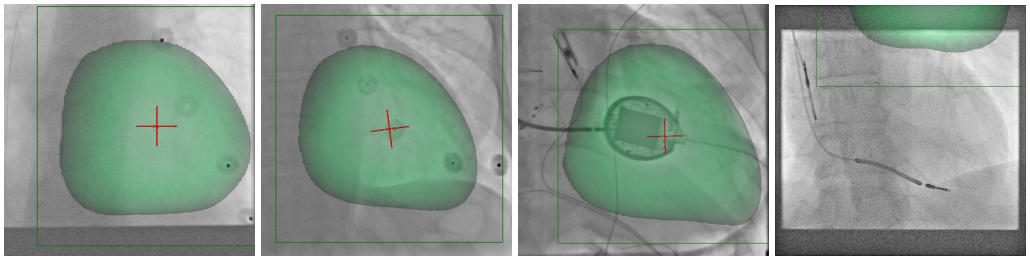
The trajectories of the agent are visualized for two highly robust cases in Figures 5.8 (a) and 5.8 (b), a robust case in Figure 5.8 (c), and the case showing the lowest robustness in Figure 5.8 (d). High robustness was shown in 15 patients, where the agents paths have always converged to the same position. In one case, the final positions were in a less confined area, see Figure 5.8 (c). In two cases the agent has only diverged for a few starting positions. In the case shown in Figure 5.8 (d), some paths are diverging (the agent has left the image), and the area, where most paths end, is not well constrained.

It is to be noted that the images were acquired in the standard clinical workflow, thus they have different acquisition parameters. This results in highly varying properties, such as image quality, FOV, or resolution. Additionally, artificial objects are often in the FOV, such as fiducial markers (Figures 5.8 (a) and 5.8 (b)) or interventional devices, such as catheters or even an ultrasound transducer, see Figure 5.8 (c). The registration appears to be robust against most factors, such as fiducial markers or FOV. Cases involving multiple devices (catheters, leads) or devices of larger extent (ultrasound transducer) are more challenging. The robustness is generally lower in these cases. The case of lowest robustness (case 15) has the lowest signal to noise ratio and implanted electrodes. Furthermore, substantial collimation, covering parts of the heart shadow is present. Initializations in proximity of the collimation result in multiple divergent trajectories. The implanted devices and the collimation appear to pose the main limitation in performance.

The accuracy was evaluated visually for randomly sampled results, showing the LV model overlayed on the X-ray images, see Figures 5.8 (e) to 5.8 (h). In cases showing robustness, the border of the overlayed LV model is aligned well with the LV shadow in the X-rays.



(a) C8: Highly robust (b) C12: Highly robust (c) C16: Robust (d) C15: Least robust



(e) C8: Success (f) C12: Success (g) C16: Success (h) C15: Failure

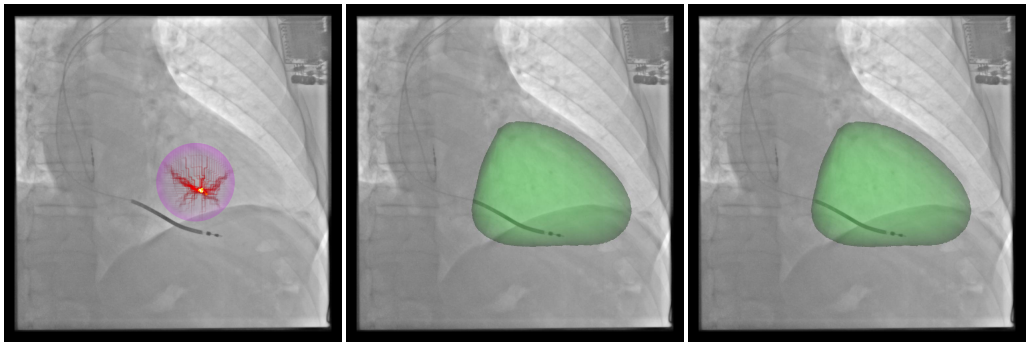
Figure 5.8: Cases showing different degrees of robustness on the AP range. (a-d) Convergence of the center point through the agents actions from various starting positions on the boundary of the purple circle. (e-h) Randomly chosen exemplary results.

RAO range

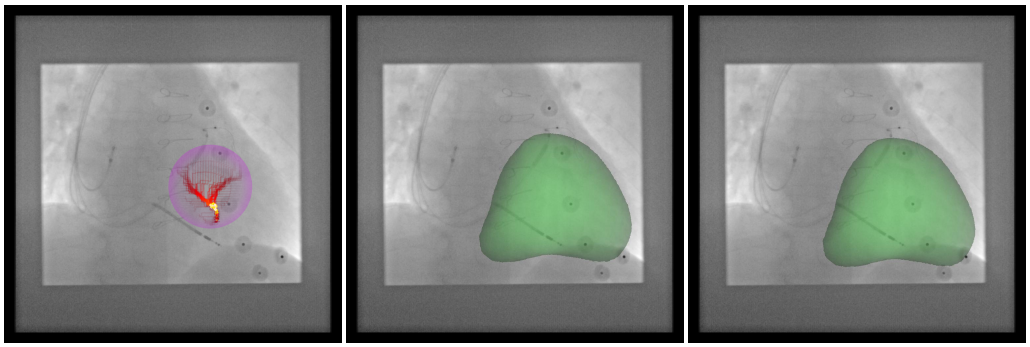
The registration was successful for the majority of perturbations in every case. The variation of the final positions was minimal. Exemplary trajectory covergence plots are shown for two cases in Figures 5.9 (a) and 5.9 (d). The trajectories of the central landmarks converge to the same point, resulting in approximately the same registration matrix for any starting perturbation. Randomly selected sample model-X-ray overlay results can be seen in Figures 5.9 (b), 5.9 (c), 5.9 (e) and 5.9 (f).

LAO range

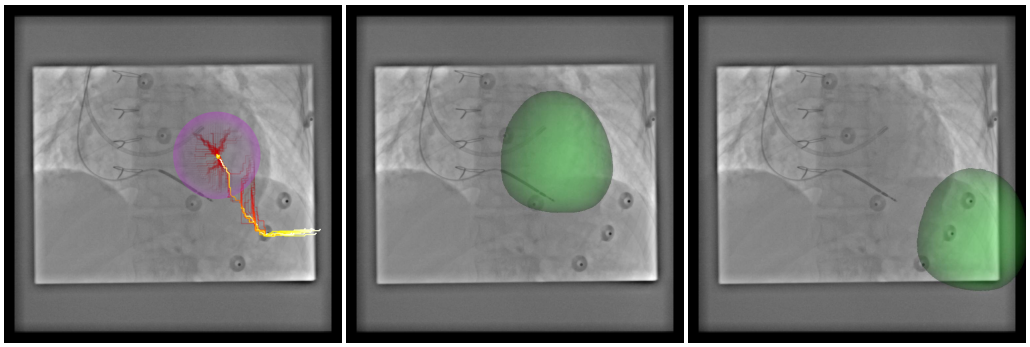
The qualitative convergence results on the LAO range were significantly worse than those on the AP and RAO ranges. There is an abundance of diverging trajectories in the majority of the cases. Similarly to the AP range, trajectories in close proximitiy of collimation edges tend to diverge more. All paths diverge in two cases and there are two cases in that all paths converge. In the remaining six



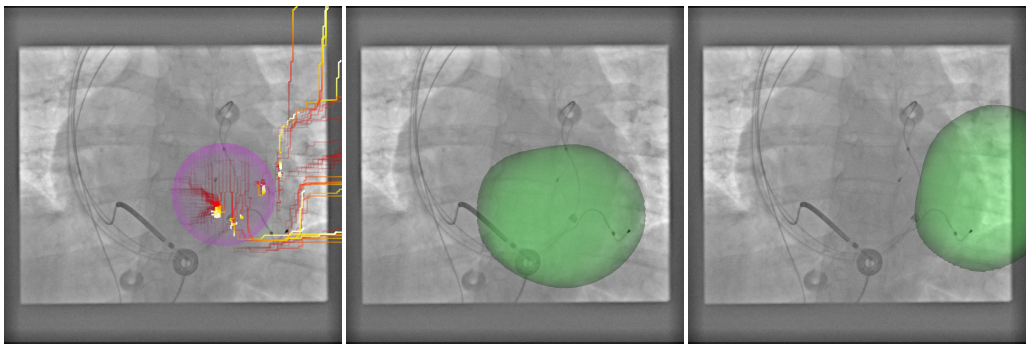
(a) RAO – C1: Highly robust (b) RAO – C1: Success (c) RAO – C1: Success



(d) RAO – C6: Highly robust (e) RAO – C6: Success (f) RAO – C6: Success



(g) LAO – C1: Limitedly robust (h) LAO – C1: Success (i) LAO – C1: Failure



(j) LAO – C9: Not robust (k) LAO – C9: Success (l) LAO – C9: Failure

Figure 5.9: Case samples to demonstrate robustness on the RAO and LAO ranges. (a, d, g, j) Agent's trajectories from different starting perturbations (perimeter of purple circle) in four clinical cases and sample overlay images.

cases, most of the trajectories converge, however, there are numerous diverging trajectories. Two such convergence visualizations are illustrated in Figures 5.9 (g) and 5.9 (j). The overlay images show a good alignment in the majority of the converging cases, see Figures 5.9 (h) and 5.9 (k). In diverging cases, the agent moves the mesh model out of the X-ray image borders, see Figures 5.9 (i) and 5.9 (l).

5.3.2.2 Quantitative Evaluation

To measure robustness, the variance of the final registration state for different perturbations was observed. The models were perturbed on a regular grid of translations (-30 mm to 30 mm, with 5 mm sampling) with random rotations ($-15°$ to $15°$), starting from a rough initial alignment, resulting in 169 perturbations per case. This was performed for all three standard CRT angulations. The models were re-registered to the X-ray images. The median final position $\tilde{\boldsymbol{x}}_f$ of the cross landmark was used as a reference. The L2 norms of the final positions $\boldsymbol{x}_f$ were computed relative to this position for each dataset separately:

$$e_f = \|\boldsymbol{x}_f - \tilde{\boldsymbol{x}}_f\|_2. \tag{5.4}$$

AP range

The resulting deviations $\boldsymbol{e}_f$ show minor variance, see Figure 5.10. In some patients, such as patient 15, the trajectory was diverging for perturbations at the edge of the capture range. In ten patients there was no divergence. The worst patient (15) had 14 % outliers. It is to be noted that errors above the training range (35 mm) are diverging trajectories. The median deviation was approximately 1 mm in every case. It was below 5 mm in 97.16 % of all cases and 90 % of all deviations were below 1.42 mm.

RAO range

The results on the RAO range were highly similar to those on the AP range, see Figure 5.11. The registration is highly robust throughout the cases. The

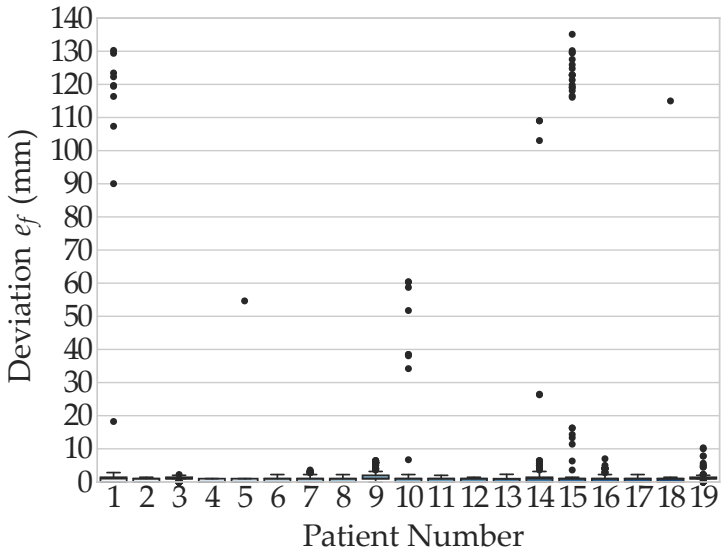


Figure 5.10: Deviations of results from the median for the AP angulation range. The points mark the outliers.

distributions of deviations from the median final registration position show only minor variance. The number of outliers is minimal 2.26 %. The median deviation was, similarly to the AP range, approximately 1 mm in every case. It was below 5 mm in 94.87 % of all cases and 90 % of all deviations were below 3.6 mm.

LAO range

On the LAO range, however, the distributions show significantly lower robustness, compared to the AP and RAO ranges, see Figure 5.12. Seven cases have a median deviation of approximately 1 mm, however, the distributions have a small variance in only half of the total cases (1, 3, 4, 5, 8). Additionally, the registration results in outliers in all cases, except for case 8.

5.4 Discussion and Conclusion

In this chapter, a machine learning-based method for registering 3D preoperative models to 2D intraoperative images for cardiac interventions, such as CRT, was presented. The method, analogously to the adjacent anatomy-based approach described in Chapter 4, is agnostic to the preoperative modality, it can be, e.g., MR, CT, or ultrasound imaging, since instead of the raw image data, 3D models are

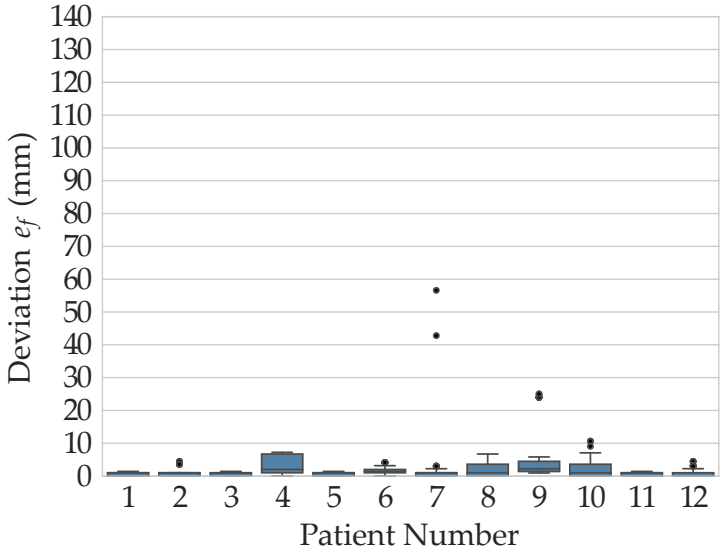


Figure 5.11: Deviations of results from the median for the RAO angulation range. The points mark the outliers.

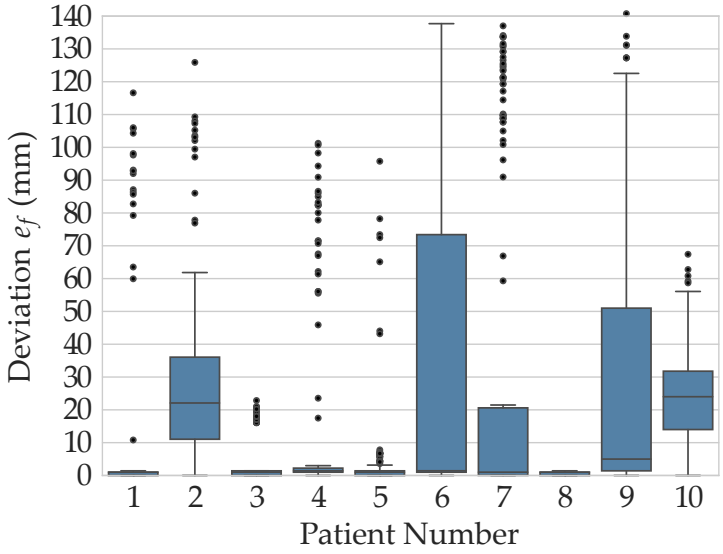


Figure 5.12: Deviations of results from the median for the LAO angulation range. The points mark the outliers.

registered. The models are often available from the standard clinical work. The method registers to a single, non-contrasted X-ray image, integrating seamlessly into the standard clinical workflow.

To register preoperative models, i.e., the LV, to X-ray fluoroscopy, imitation learning was performed. Due to the limited availability of labelled data on the target domain (MR and X-ray data pairs), the method was trained purely on synthetic data. The synthetic training data consisted of 1) mesh models extracted from CT data by automatic segmentation and 2) artificially generated X-ray images, DRRs projected from CT data. The benefit of using only CT data in training, besides being available in large quantities, is that both the 3D and the 2D data originate from the same CT volumes, thus the GT alignment is given and is error-free.

The available 802 CT volumes were partitioned into 702 training and 100 evaluation datasets. Three ANNs were trained, one for each of the three possible CRT angulation ranges: AP, RAO, and LAO. First, it was proven, that an individual network for each angulation range is more performant than training a single network on the whole possible angulation range.

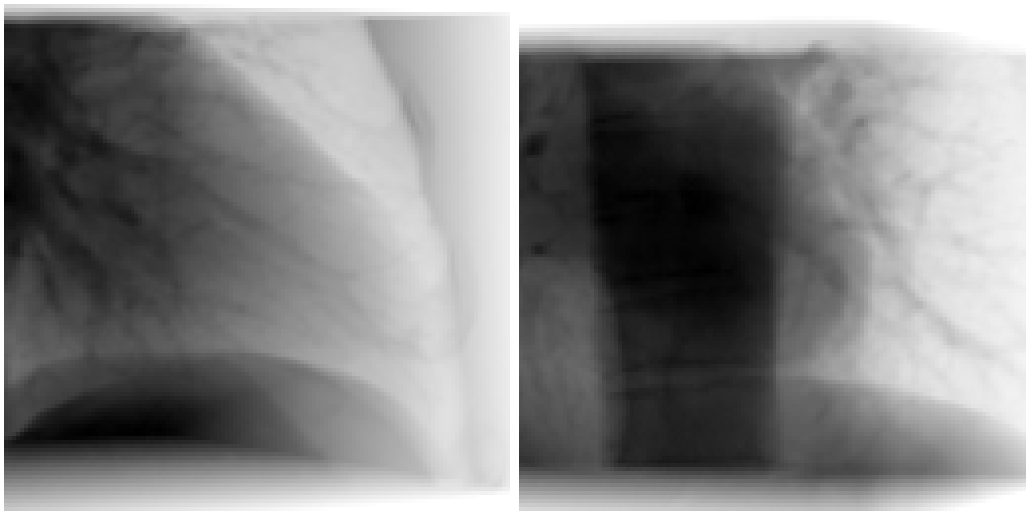
The trained networks were evaluated extensively qualitatively and quantitatively on synthetic and clinically acquired CRT data. The synthetic data experiments have demonstrated high accuracy on all three angulation ranges. The imitation learning agent's performance was superior to manual and gradient-based approaches. The registration accuracy was highest on the LAO and lowest on the RAO angulation range. This can be explained by that in RAO images the heart shadow is often not comprised of only the LV, the other chambers of the heart, i.e., the RV, contribute as well. Thus, the strongest feature, the heart shadow in the DRR images, does not correspond to the border of the LV in such cases.

In the clinically acquired CRT cases, MR to X-ray data was registered. It was demonstrated that the system trained fully on synthetic data generated only from CT volumes, can register models from MR to X-ray images. Due to a GT

registration not being available, perturbations were applied to a rough manual registration of the LV model. The perturbed meshes were registered and their final positions compared. In the AP and RAO ranges, almost all perturbations were re-registered to the same confined area, demonstrating good robustness, showing good transfer from the synthetic training domain to the clinical target domain. However, outliers were also present, especially for larger perturbations that are close to the edge of the capture range of the approach.

The robustness was significantly lower on the LAO range. This is uncorrelated with the accuracy results on the synthetic data, where the LAO range has shown the highest accuracy. This can be explained by the LAO network overfitting to the synthetic results the most, providing lower performance on the clinical CRT data. The main reason for this could be that the LAO DRRs appear more dissimilar to their X-ray counterparts than the AP and RAO DRR images. The higher dissimilarity of LAO DRRs results from the more prominent appearence of the aorta than on the other angulation ranges. The CT volumes were acquired with contrast agent injections. The chambers of the heart and great vessels have significantly improved contrast compared to standard, non-contrast-enhanced CT volumes. One of the most prominently enhanced organs is the aorta. In RAO projections it is not visible, in AP rarely, however, in LAO projections it is often overlaying the LV, see Figure 5.13. The aorta is not distinguishable in clinical X-rays, however, it is one of the most prominent structures in the generated training DRRs. The imitation learning agent learns different features from the training images than extracts from the clinical target images, resulting in less successful transfer to the target domain of clinical X-rays.

A further factor contributing to the lower performance on the LAO range is that the heart shadow is mainly comprised of the LV's projection. This can simplify the learning process and the imitation learning agent learns to focus mainly on the strong edge features of the heart shadow. On the other ranges, especially on RAO, multiple chambers can contribute to the silhouette of the heart in the X-ray images. Thus, it is not sufficient to focus on the heart shadow, more



(a) RAO (b) LAO

Figure 5.13: Examplary samples of generated DRRs of a patient for the (a) RAO and the (b) LAO angulation ranges. The main difference between the DRRs is that the shadow of the aorta overlays the LV in LAO images (dark band in the center of the LAO projection). Thus, the RAO images are more similar to clinical X-rays than the images of the LAO images.

subtle features need to be learned. The need to use more and weaker features can prevent overfitting.

The third contributing factor on the LAO range was the proximity of the initialization to the collimation edges. This has resulted in registration failures for a small number of initializations on the AP range. The imitation learning agent was not trained with collimated images, thus it is not able to provide robustness in highly collimated images consistently.

This chapter has demonstrated that an imitation learning agent, trained only on synthetic data, is capable of registering an LV mesh model extracted from MR to X-ray images with high robustness. In interventional procedures, highest degrees of robustness are desireable. Additionally, performance after transferring a system to the target domain that was trained only on synthetic data can vary. In the current setup, no explicit transfer learning is performed. The performance on clinical data can increase through transfer learning. The next chapter introduces the concept of domain randomization to increase transfer performance, providing consistent robustness, without explicit optimization of

data generation or training parameters.

Chapter 6

Domain Randomization in Learning-based Registration

In Chapter 5, an imitation learning agent was described that is capable of registering a mesh model of the left ventricle (LV) to an uncontrasted X-ray image. The agent was trained purely on synthetic data, digitally reconstructed radiographs (DRRs). It was evaluated qualitatively and quantitatively on the same domain and on the domain of clinically acquired X-ray data of multiple angulations, i.e., anterior-posterior (AP), right anterior oblique (RAO), and left anterior oblique (LAO), from cardiac resynchronization therapy (CRT) procedures. However, no explicit optimization of system components, such as DRR generation parameters, was performed, the imitation learning agent has shown high accuracy and robustness. This chapter 1) explores the influence of training parameters on the transfer learning performance from synthetic to clinical data by comprehensive evaluation and 2) proposes a method based on domain randomization (DR) to improve the reliability of the transfer. The number of training datasets is increased significantly, to investigate the influence of the amount of training data on transfer performance.

6.1 Introduction

In the previous chapter, an imitation learning-based system for registering 3D preoperative data to 2D X-ray images has demonstrated the feasibility and benefits

of training on fully synthetic data. The system generated synthetic images from computed tomography (CT) volumes, thus knowing the ground truth (GT) transformations. These synthetic X-ray images are commonly referred to as DRRs [147]. The registration system trained on DRR data has shown high accuracy on the domain of the training data (DRRs) and acceptable performance on the target domain of clinical X-ray data. However, transfer performance is highly affected by training parameters, such as the weight initialization of the neural network or the training data order: different local minima might be reached, resulting in highly similar performance on the synthetic training domain, but highly variable results on the target domain of clinical X-ray data, see Figure 6.1. DRR generation algorithms do not model every aspect of the X-ray formation. This causes images to appear differently than clinically acquired X-rays. Algorithms used for image guidance have to be extremely robust against perturbations. In the case of the X-ray data, these could be intensity variations, due to different dose settings, variation of the collimation and shutters, or devices in the field of view. These variations can appear differently, or may not be present, in the artificially generated DRRs. The transfer, thus the robustness of the registration can suffer. The perfomance reduction through transferring to the target domain is called the reality gap.

The principle of differences between feature sets of DRR and X-ray data can be illustrated with the help of a Venn diagram, as depicted in Figure 6.2. The set of DRR features has a significant intersection with the set of X-ray features. The intersection, illustrating the shared features between DRRs and X-rays, enables a system that was trained purely on the DRR domain, to perform adequately on X-ray images. If the sets are disjoint, the trained system will not transfer, because the X-ray data does not share any features with the DRR data. The registration will fail on X-ray data, corresponding to a large reality gap. On the other hand, the larger the intersection, the higher the transfer performance will be.

Several approaches have addressed the problem of bridging the reality gap, including Heimann et al. [148], where a classifier is adapted to the target domain,

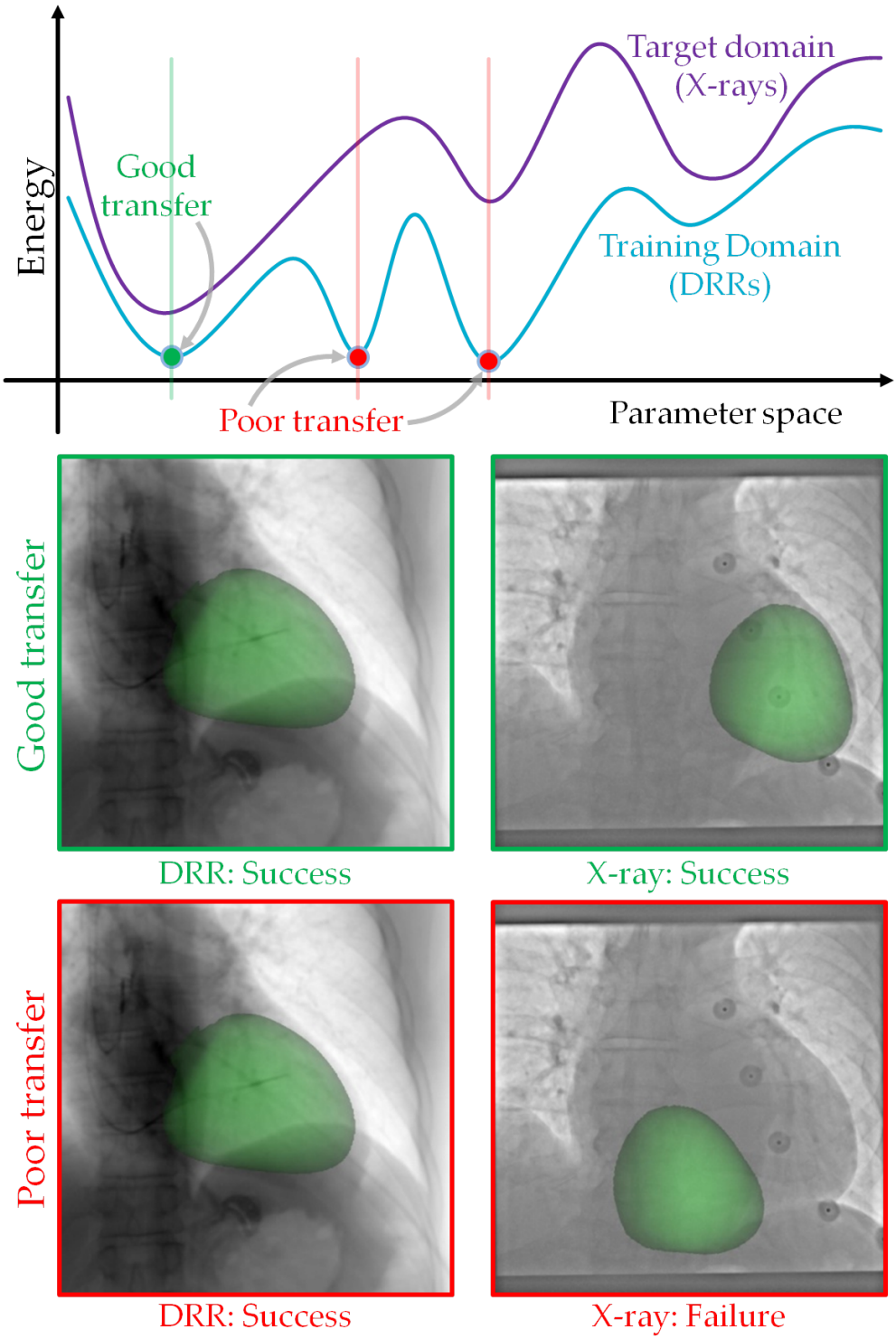


Figure 6.1: Domain transfer performance can be parameter-dependent. Training the agent on DRRs (blue) can result in highly similar local minima for different training parameters. The found minima may not correspond to minima of the loss on the target domain (X-rays) (purple). One minimum (green) of the DRR energy landscape corresponds to a minimum of the target domain's landscape. Transfer performance will be high for this minimum: optimal results on both domains. However, low for other minima (red): high performance only on DRRs.

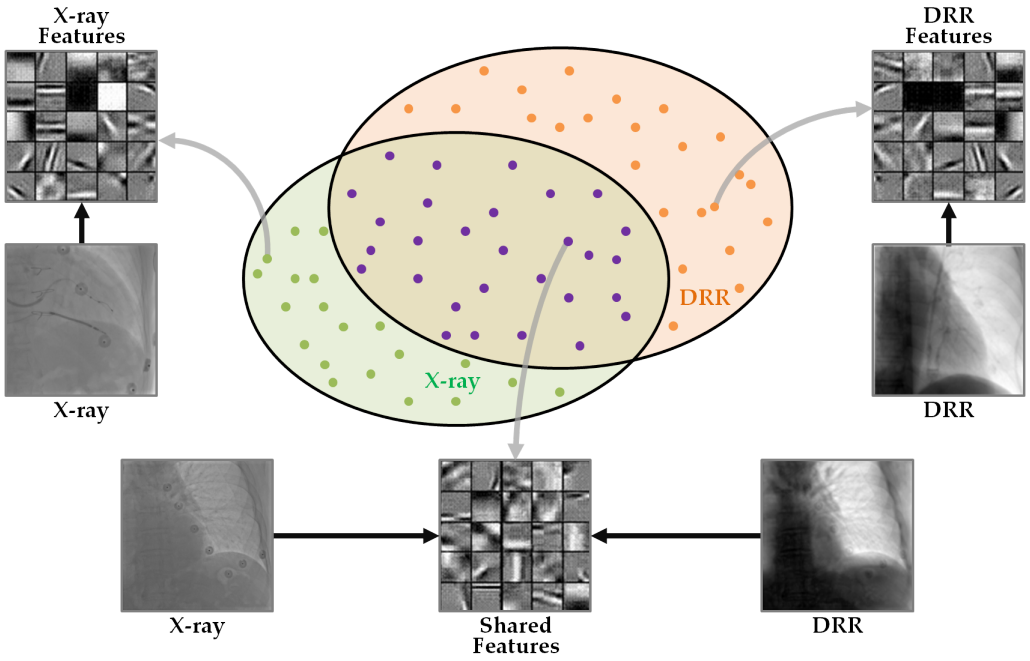


Figure 6.2: The feature spaces of the DRR and the X-ray domains represented by a Venn diagram. The generated DRRs distribution does not model all aspects of the X-ray distribution, however, the distributions overlap. A system trained on the DRR domain will transfer to the X-ray domain, if there is an intersection of the two sets. The larger the intersection, the more successful the transfer can be.

without requiring labeled data on the target domain. A set of classifiers is trained to identify the pose of a transesophageal echocardiography (TEE) probe in X-ray acquisitions. The classifiers are trained through marginal space learning [121] and implemented as probabilistic boosting trees [149]. Due to limited amounts of (labeled) training data being available, training is performed on labeled, synthetically generated X-rays (additive combination of TEE probe DRRs and background X-ray images) and unlabeled X-rays of the target domain. Instance weighting is performed to realize domain adaptation, the transfer between the synthetic and the clinical target domains. The weights are learned by training a classifier to distinguish between synthetic and clinical data. However, it could fail, because instance weighting will not be sufficient, if the synthetic and cinical training data distributions are far from each other, e.g., if the synthetic images are not of sufficient quality. The approach would also fail, if the training and target clinical distributions are not sufficiently similar, e.g., the data are acquired at

different clinical sites or with different imaging devices. The approach needs to be retrained for such newly acquired datasets.

Domain adaptation can also be performed through task-driven generative adversarial networks (GANs) [150]. A cycle consistent GAN (Cycle-GAN) [151] can be trained in conjunction with a segmentation (task) network to perform unsupervised segmentation. The segmentation network is pretrained on synthetic data and frozen, however, its loss is backpropagated through one arm of the Cycle-GAN. The segmentation network guides the image generation process of the Cycle-GAN, to generate synthetic images that maximize the performance of the segmentation network. The method can significantly increase the perfomance of the segmentation network, but cannot achieve the results of fully supervised training. Although training can be performed in a fully unsupervised manner, GANs can be difficult to train, can become unstable, and can show mode collapse. Furthermore, the method may not transfer to unseen X-ray images, if the unlabeled X-ray data come from a different distribution.

Advanced simulation methods can be used to generate higher quality DRRs that appear more similar to the target domain's images, such as Unberath et al. [152]. The approach decomposes the CT through a convolutional neural network (CNN) into different tissue types, i.e., air, soft tissue, and bone. An attenuation image is composed of individual forward projections of the tissue types. X-ray scatter is approximated through another CNN trained on Monte Carlo simulations. Quantum noise and readout noise are injected through simple Poisson and Gaussian models. The method is shown to generate realistic X-rays, though a landmark detection algorithm demonstating high performance on the synthetic data. However, the method is configured on a limited set of training data, thus might overfit to it. Additionally, the approach consists of multiple components that are approximations of the physical processes. The errors of the individual approximations can accumulate, generating potentially inaccurate simulated images. Furthermore, the material decomposition handels only three tissue types, unexpected materials, e.g. metal of implants, will not be accurately

simulated.

The collective weakness of these approaches is that they adapt/fit their models or features to samples of the target domain. If the samples are not representative of the target domain, the adaptation will overfit to the available samples. If such a system has to process data from a new site, device, or of unexpected imaging parameters, performance can decrease significantly.

This chapter addresses the problem of overfitting to the training, or a non-representative target domain, by using DR [153]. In DR the training data is augmented with unrealistic transformations. The transformations introduce such a large variation to the data that the target domain will appear as just another variation to the learning system. This approach was shown to help transfer to the target domain in non-medical applications, e.g., autonomous driving [154] or robotic control [155].

In this chapter, the usability of DR in the medical imaging context is demonstrated. It is utilized, to enhance the transfer performance of the imitation learning agent-based system, described in Chapter 5. It is shown that, if trained only on artificially generated data (DRRs), DR can greatly increase robustness against geometric perturbations, compared to the standard imitation learning-based approach, without explicit transfer learning. Furthermore, it is demonstrated that, without DR, variation of training parameters, such as network weight initialization or training data order, can result in inconsistent results on the target domain. Performance with DR, however, appears to be agnostic to variation in training parameters, thus produces substantially more consistent results.

6.2 Methods

The registration framework with domain randomization is illustrated in Figure 6.3. The imitation learning-based model-to-X-ray registration system from Chapter 5 is used for 3D/2D cardiac registration. It is trained with purely CT volumes, without any real X-rays from the target domain. The CT volumes are used for generating two 2D images: 1) DRRs are projected and 2) the LV mesh

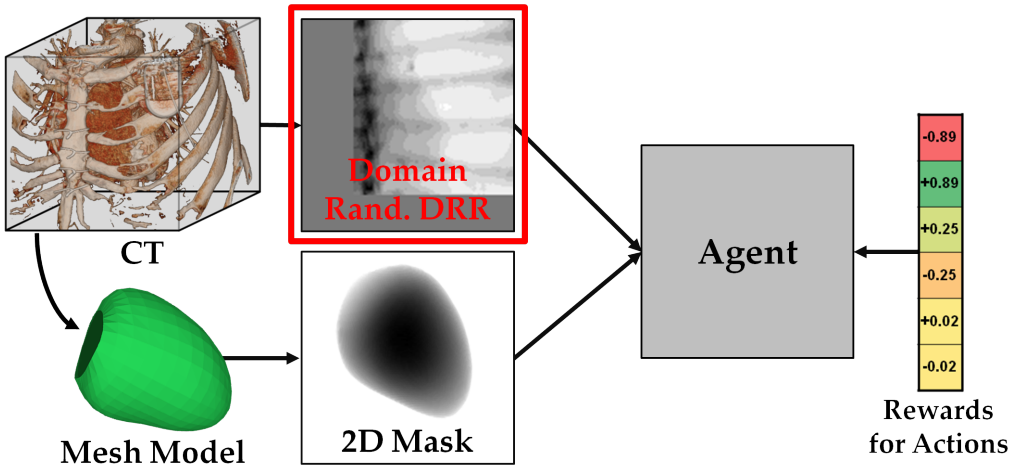
model is segmented, perturbed in 3D and projected into 2D. The main difference compared to the system described in Chapter 5 is that the DRR images are domain randomized. Since both images are generated, the ground truth transformation is known between them. This is used to compute rewards for each possible action. The rewards and the two 2D images (domain randomized DRR and mask image) are used to train an artificial agent represented by a CNN, see Figure 6.3 (a). For inference, the artificial agent is being shown the two images and is able to predict the reward for each possible action. The action with the highest reward is chosen and is applied to the 3D model. For an overview see Figure 6.3 (b).

In the following sections, 1) the transfer learning considerations regarding the involved DRRs and X-rays are highlighted and 2) the principle of DR and how it is applied to generate the DR DRR images is described in detail.

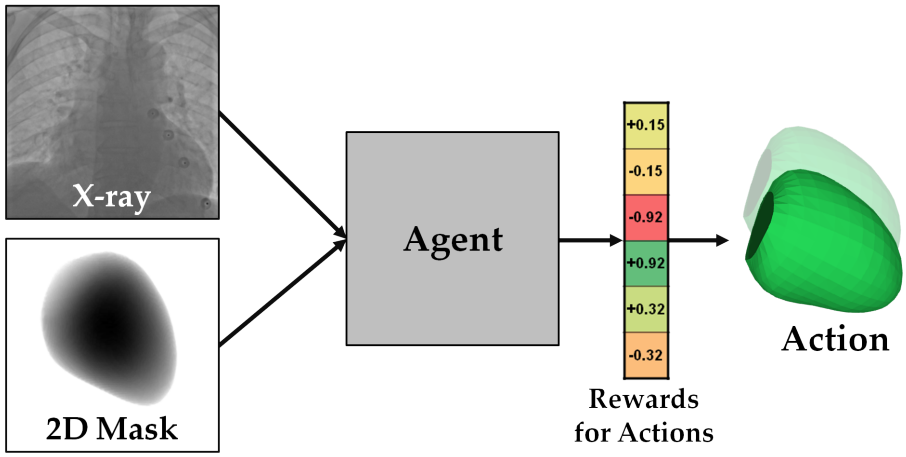
6.2.1 Transfer Learning

As described in Chapter 5, the training of the artificial agent is performed on fully synthetic data. The agent has shown favorable performance in terms of accuracy on the synthetic domain (DRRs) and robustness on the target domain (clinical X-rays). However, transfer to the target domain can be unreliable, results might not be reproducible. How well the imitation learning agent transfers to real data, is influenced by three main factors: 1) the appearence of the training data (similarity to target data), 2) the weight initialization of the neural network (NN), and 3) the training data order.

The similarity between the training domain and the target domain is the main factor in performance after transfer. If the training DRRs are sufficiently similar to the target X-ray images, the network will transfer with similar performance. This can be achieved to a certain degree by optimizing DRR generation parameters, but it is highly challenging. If the sample images are not representative of the target domain, the network will not generalize well and might not perform to the same standard on the newly acquired data as on the training domain. In clinical applications, images can come from different scanners, i.e., models or



(a) Training workflow



(b) Evaluation workflow

Figure 6.3: Overview of the imitation learning-based registration framework workflows enhanced with DR. (a) In the training workflow, DR DRR images are generated, instead of standard DRRs. (b) No amendments were implemented in the evaluation workflow of the imitation learning agent described in Chapter 5.

manufacturers, can be acquired by different operators, and can come from different sites. To compensate for the performance reduction due to the differences of the previously unseen data, the majority of current machine learning-based systems have to be retrained on or finetuned to the newly acquired data.

The second factor is the initialization of the weights in the NN representing the agent. The initial weights define the starting position in the optimization space of training. From different starting positions, different local minima can be reached that can result in highly similar results on the training, but inconsistent results on the target domain.

The effects of the variation of the training data order can have similar influence. If the same weights are used for initialization, but the network sees the training data in a different order, the optimization can take different steps and converge to different local minima, resulting in different results on the target domain.

6.2.2 Domain Randomization

Domain randomization (DR) [153] is an approach to transfer a deep NN trained on only simulated images to the real world. Instead of high quality simulation images, DR relies on the variability of the low quality simulation images. The main idea is to randomize the simulator and generate a large variation of simulated images, such that the real images become just another variation. Thus, if the randomization of the image generation is performed adequately, the transfer to the target domain will be more efficient and performance on the target domain will increase.

For the present problem of imitation learning-based cardiac registration, instead of using a realistic X-ray simulator, it is proposed to use a simple ray tracing-based DRR renderer. The concept can be illustrated by extending the Venn diagram presented in Section 6.1, see Figure 6.4. The sets of X-ray features and DRR features have a substantial intersection, because the method transfers (with limited performance) to the X-ray domain, even if trained exclusively on DRR

data, without seeing clinically acquired images. If there were no shared features, there was no intersection of the sets of X-rays and DRRs, the approach would not transfer at all. To ensure a more successful transfer, the DRR generation is randomized that the resulting features cover a large portion of the whole feature space (blue dots in Figure 6.4). If the DR images cover most of the feature space, the X-ray images will be just other variations from the sampled whole feature space, thus the approach will generalize to the X-rays too, even if no X-ray similar images are generated.

The main differences between the generated DRRs and clinical X-ray data are in the 1) image contrast and brightness, 2) noise, and 3) collimation. The differences in contrast and brightness are clearly visible and measureable through, e.g, histogram analysis of the intensity values. The image intensities can be refined to be more similar to the X-ray intensities, through optimizing the DRR rendering parameters of the low quality simulator, such as Hounsfield unit (HU) thresholds. However, the intensity transfer function is linear, it cannot model non-linearities of real X-ray projection accurately, such as scattering or beam hardening. Furthermore, the training DRRs are generated from contrast-enhanced CT volumes. The contribution of contrast-enhanced structures to the projection, such as the aorta or the cardiac chambers, as described in Section 5.4, can result in highly dissimilar DRRs to clinical X-rays. This effect was most significantly observable on the LAO angulation range of projections.

To account for these differences, the HU values of the CT volumes are randomized. Noise is also not modeled by the simple simulator. The experimental outcomes in Chapter 5 were not dependent on noise. The clinical X-rays in the evaluation were acquired with highly varying parameters. Several were acquired with low dose fluoroscopy, but there was no correlation between dose and registration performance. However, X-ray collimation has a more significant role. It was shown in Chapter 5, if registration is initialized close to the collimator edge, the imitation learning agent might fail.

To increase the generalization of the imitation learning agent, the DRR

generation is performed with randomized 1) HU values of the 3D CT data and 2) a border representing collimation added to the 2D DRR projections. Before computing the DRR ray tracing, an intensity mapping to the HU values of the CT data was applied either globally or locally. For global randomization, a non-linear function was applied to transform the voxel values.

The chosen function was the cumulative distribution function (CDF) of a beta distribution, the regularized incomplete beta function:

$$I_x(\alpha, \beta) = \frac{B(x; \alpha, \beta)}{B(\alpha, \beta)}, \tag{6.1}$$

with

$$B(\alpha, \beta) = B(1; \alpha, \beta), \tag{6.2}$$

where

$$B(x; \alpha, \beta) = \int_0^x t^{\alpha-1}(1-t)^{\beta-1} dt \tag{6.3}$$

and the parameters α and β are randomly sampled from a uniform distribution of $[0.5; 5.0]$ and x is the voxel intensity to be transformed. The CDF of the beta distribution provides a smooth intensity transfer function between 0 and 1 (the minimum and maximum of normalized intensity values), while requiring only two parameters.

For the local randomization, the intensity range of the CT data was subdivided into 5 to 15 non-overlapping intervals. The start and end of the interval ranges was defined by evenly spacing the whole range and applying a random perturbation to the start and end values, sampled from a Gaussian distribution (mean = 80 and standard deviation = 40). For each intensity range, one of three randomization options was chosen: 1) the non-linear mapping described above; 2) inversion of the intensity range; 3) shift of the intensity values by adding a random value of a uniform distribution on the interval $[-100; 100]$. For a sample of the resulting images see Figure 6.5 (c). For a comparison, a sample of standard, generated DRR images is shown in Figure 6.5 (b).

The acquired X-ray images of the target domain are often collimated. Thus,

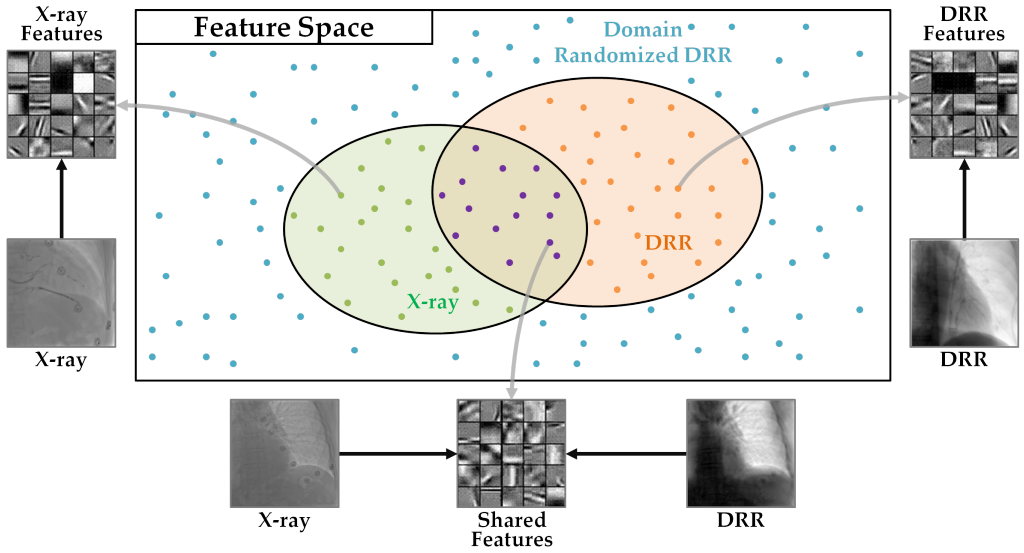


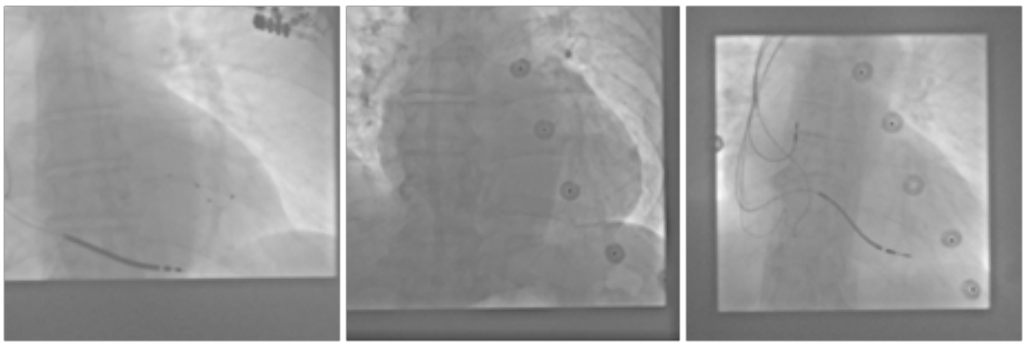
Figure 6.4: The concept of DR in feature space. Every point represents a sample (a set of features) in the respective domain (green: X-ray, orange: DRR, purple: shared features, blue: DR DRR).

in addition to the intensity and imaging geometry randomizations, bands of random intensity were added to the generated images, see Figure 6.5 (d). The intensity was varied in the range of all possible intensity values $[0; 1]$ and the size of the borders was maximally 25 % of the image size in the respective direction.

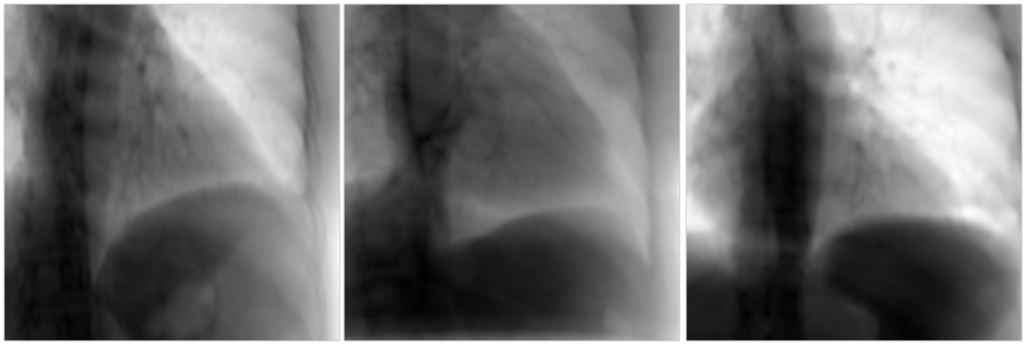
6.3 Evaluation and Results

The effects of DR are demonstrated on the image registration task with an imitation learning agent that was described in Section 6.2, to improve its transfer performance to the target domain. The DR imitation learning agent was trained only on synthetically generated data, similarly to that of Chapter 5. The number of training CT volumes was increased by 1009, to 1711 CT volumes of 799 patients, to assess the influence of training on a larger dataset on transfer performance. The evaluation was performed on the two datasets described in Section 5.3: 1) synthetic data from the same domain that the training data originates from; 2) clinical patient data, from the target domain.

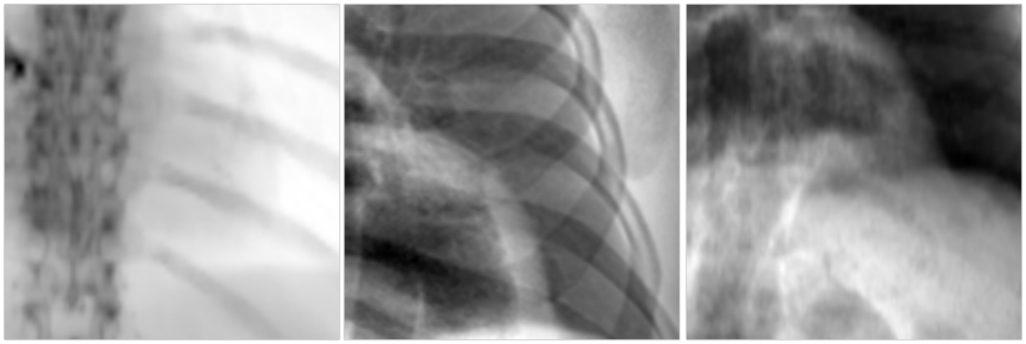
First, the effect of DR is presented for a certain set of seed values for 1) random weight initialization and 2) random data shuffling. Then the seeds are



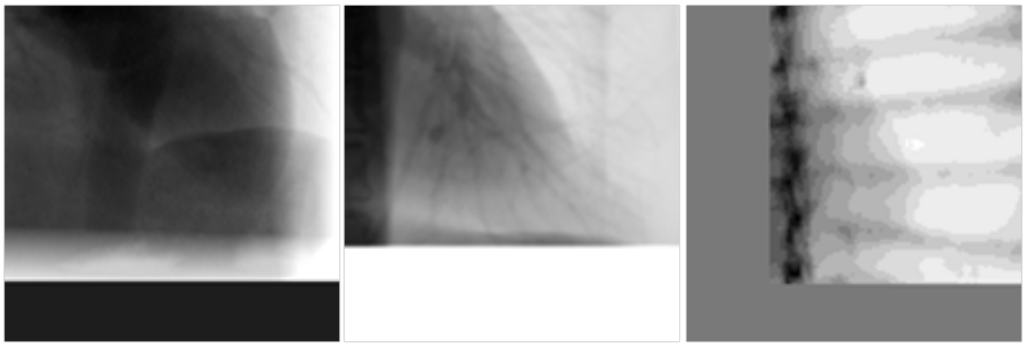
(a) Traget domain: X-rays



(b) Standard DRRs



(c) DR intensity DRRs



(d) Synthetic DR training domain: DR intensity and collimation DRRs

Figure 6.5: Sample images of domains involved in DR registration framework.

varied to demonstrate the inconsistency of performance after the transfer from the synthetic training domain of DRR images to the target domain of clinically acquired X-ray images and to show the consistency of the domain randomization results.

6.3.1 Evaluation Data

The synthetic test dataset was the same as in Chapter 5, consisting of 100 CT volumes of 100 patients. There was no overlap between training and test patients. The 3D mesh model of each patient was perturbed 10 times sampled from a uniform distribution, resulting in 1000 artificial test cases. Due to the known perturbation, the ground truth registration is available. Accuracy is measured by the L2 norm of the endpoints of a 3D landmark at the center of the LV model, as described in Section 5.3.1.3.

The number of clinical patient datasets was increased from 19 to 21 datasets. Each patient dataset has one MR volume of that the mesh model was extracted and at least one X-ray image to register to. In the two additional patient datasets, 20 and 21, X-ray images of all three angulation ranges were available. For an overview of the availability of X-ray images in the first 19 patient datasets see Table 5.4. The following experiments were performed for all three angulation ranges: AP, RAO, and LAO.

In the clinical CRT data experiments, the mesh model is perturbed, as described in Section 5.3.2.2, 169 times in a rectangular grid manner about an approximate, manual registration, in the plane of the respective X-ray image. There is no ground truth registration available for this dataset, thus the registration is evaluated for robustness, i.e., that the registration provides the same result from different perturbations, as described in Chapter 5. Accuracy is only verified qualitatively, by comparing the heart shadow in the X-rays and the edges of the projected LV mask.

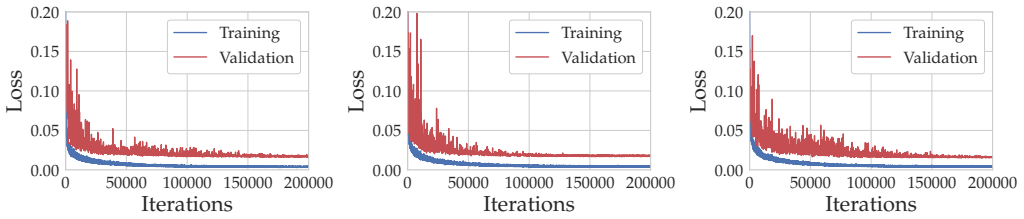
6.3.2 Baseline

The standard network (SN) without performing DR (as presented in Chapter 5) was trained with a selection of seed values for data order (D_1) and weight initialization (W_1) to provide reproducible baseline results to compare against. As described in Chapter 5, training was performed for each possible CRT angulation range, i.e., AP, RAO, and LAO, individually, resulting in three specialized networks. The number of training iterations was doubled to 200 000, to ensure convergence with the larger training dataset, approximately double the size of that in Chapter 5. The training was performed with the same parameters on the DR data, ensuring that the training and evaluation results are reproducible and directly comparable. The whole training process for the baseline seeds resulted in six networks in total, one SN and one domain randomized network (DRN) for the three angulation ranges respectively.

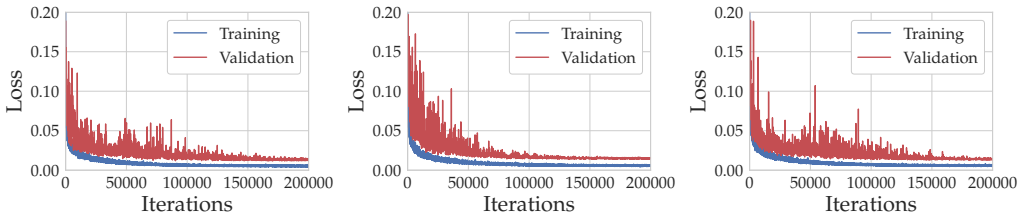
No major difference can be observed in the training curves, but the variation of the validation curve is substantially larger for the DRNs than for the SNs in all three angulation ranges. Convergence of the validation curves is minimally slower for the DRNs than for the SNs, because the variation in the training data increases through DR, thus there is more information to process and learn. The training curves are closer to each other for the DRNs after convergence, see Figure 6.6, suggesting less overfitting to the training data. The evaluation results are described and discussed individually for the three angulation ranges.

6.3.2.1 AP Range

In the AP range, the accuracy on the synthetic data with the SN is marginally higher than in Chapter 5, see Table 6.1. However, the clinical data results are minimally worse, see Figure 6.7 (a). This is due to training on significantly more synthetic data, causing the network to overfit more to the synthetic domain. The DRN has the same architecture and was trained with the same parameters as the SN. On synthetic data, as expected, the results improve only minimally, because no transfer is required.



(a) AP – SN; W_1, D_1 (b) RAO – SN; W_1, D_1 (c) LAO – SN; W_1, D_1



(d) AP – DRN; W_1, D_1 (e) RAO – DRN; W_1, D_1 (f) LAO – DRN; W_1, D_1

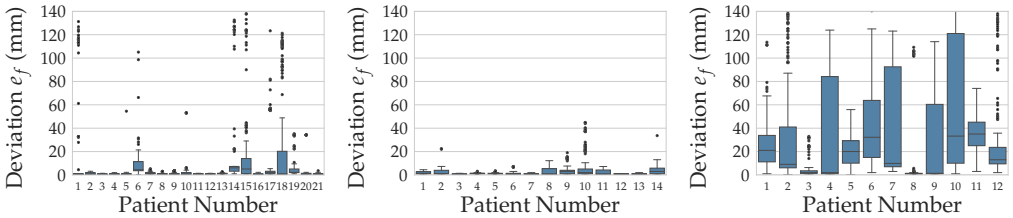
Figure 6.6: Baseline training and validation curves of the standard (SN) and the domain randomized network (DRN), with weight initialization and data order seeds, W_1 and D_1 respectively, for the three angulation ranges: AP, RAO, and LAO.

The clinical data experiment has shown, however, that the distributions are compressed and the number of outliers is greatly reduced, see Figures 6.7 (a) and 6.7 (d). Most individual cases have improved greatly in the final alignment, for samples of overlay images see Figures 6.8 (a) and 6.8 (b). The network has learned to handle intensity and collimation variations, although, it probably has never seen realistic variations of these.

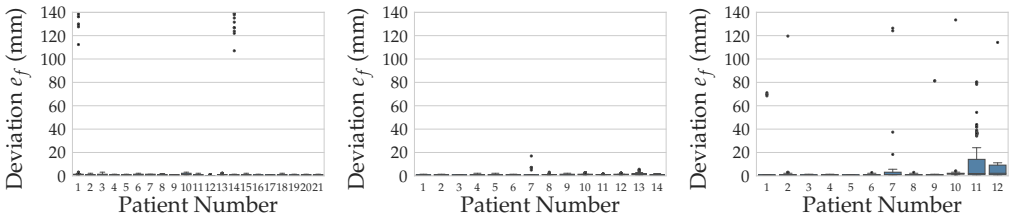
6.3.2.2 RAO Range

In the RAO range, the validation curve reaches convergence significantly earlier than in the other two ranges, i.e., AP and LAO, see Figures 6.6 (b) and 6.6 (e). This is the case for both SN and DRN. This indicates less overfitting to the training data, thus better generalization.

The registration accuracy of the DRN on the RAO range on synthetic data was marginally lower than that of the SN, see Table 6.1. Similarly to the results on the AP range, this can be explained by no domain transfer being required. The DRN does not improve registration performance on the same domain, providing virtually the same results as the SN.

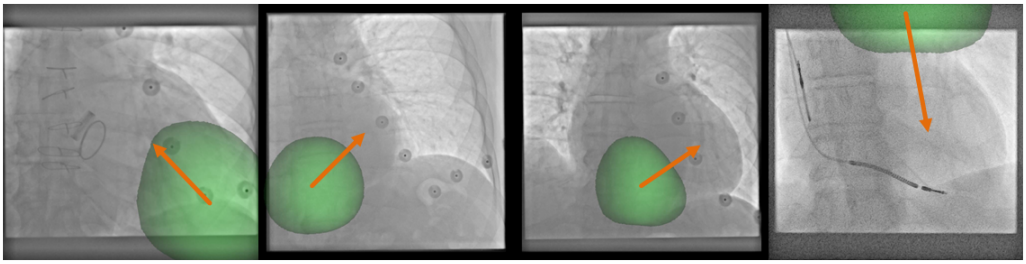


(a) AP – SN (b) RAO – SN (c) LAO – SN

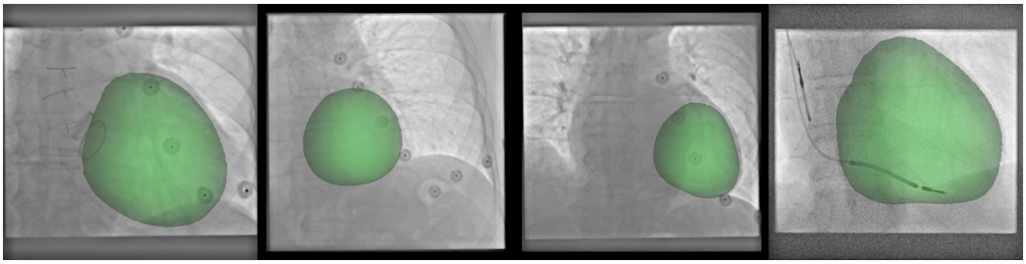


(d) AP – DRN (e) RAO – DRN (f) LAO – DRN

Figure 6.7: Effects of domain randomization on the target domain of clinically acquired patient X-ray data. Baseline quantitative results with fixed seeds (W_1 and D_1) on the three angulation ranges: AP, RAO, and LAO with the standard (SNs) and the domain randomized networks (DRNs).



(a) SN; W_1, D_1: Failed registrations



(b) DRN; W_1, D_1: Successful registrations

Figure 6.8: Effects of DR on the target domain (clinical patient X-ray data) on the AP angulation range. Baseline qualitative results with fixed seeds (W_1 and D_1) of the (a) standard (SN) and the (b) domain randomized network (DRN).

Table 6.1: Target registration error (TRE) of the cross landmark on synthetic (DRR) data initially (Start) and after registration with the standard (SNs) and the domain randomized networks (DRNs) in mm for the three angulation ranges: AP, RAO, and LAO. As anticipated, DR does not improve results significantly, because no domain transfer is necessary. Training and evaluation is performed on the same domain, i.e., synthetic DRR data.

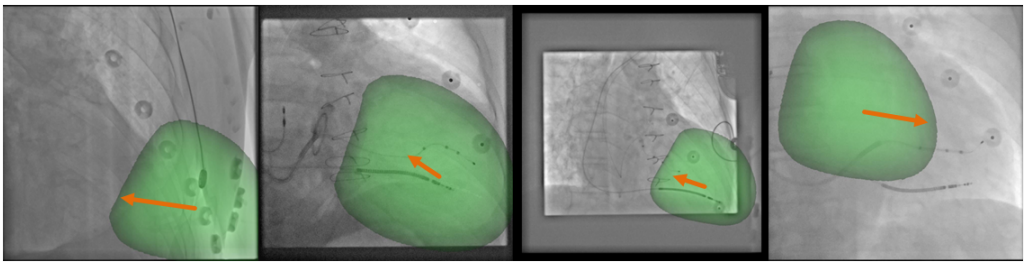
			Percentiles					
	Mean	**StD.**	50 %	60 %	70 %	80 %	90 %	100 %
AP – Start	22.41	10.60	21.27	25.47	29.34	33.25	37.12	47.33
AP – SN	2.80	2.17	2.15	2.67	3.27	4.11	5.54	13.40
AP – DRN	2.65	2.05	2.16	2.53	3.02	3.79	5.00	15.51
RAO – Start	22.71	10.70	21.36	25.74	30.17	33.67	36.86	48.45
RAO – SN	3.11	2.38	2.45	3.01	3.70	4.57	6.39	15.09
RAO – DRN	3.26	2.41	2.60	3.09	3.83	4.92	6.77	13.54
LAO – Start	22.49	10.43	21.51	25.88	29.29	32.88	36.43	47.37
LAO – SN	2.44	1.69	2.04	2.37	2.93	3.62	4.52	11.06
LAO – DRN	2.42	1.75	1.91	2.31	2.79	3.44	4.68	12.77

The experiments with the SN on the target domain of clinically acquired X-ray data have resulted in the highest robustness of the three ranges, see Figure 6.7 (b). This confirms good generalization, even without explicitly performed transfer learning to the target domain. The DRN has provided further improvement in the robustness distributions and the majority of outliers was eliminated, see Figure 6.7 (e). For a sample set of model-X-ray overlay images see Figure 6.9.

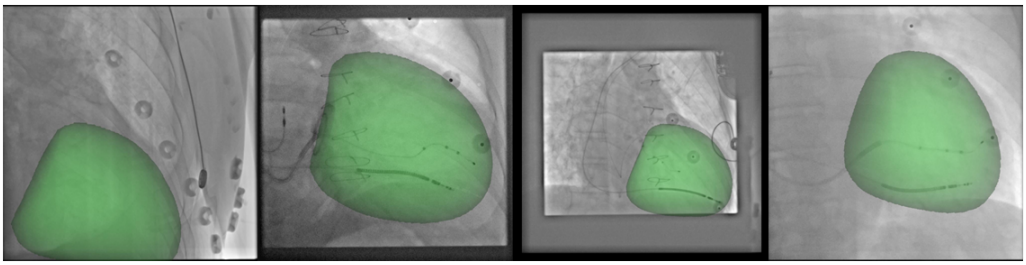
6.3.2.3 LAO Range

In the LAO range, there is virtually no difference between the TREs of the SN and the DRN on the synthetic data, see Table 6.1. This can be explained by the fact, similarly to the ranges AP and RAO, that no domain transfer is required, thus DR does not improve the results significantly.

The performance of the SN trained on the LAO range in terms of robustness was significantly lower than the SNs' on the other two ranges, i.e., AP and RAO. The results were worse than those presented in Chapter 5, because the network

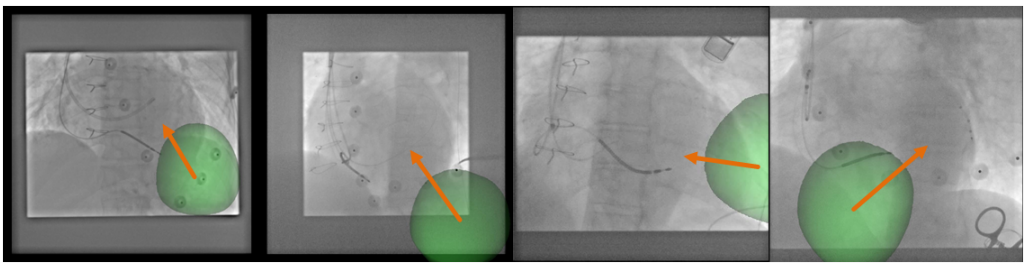


(a) SN; W_1, D_1: Failed registrations

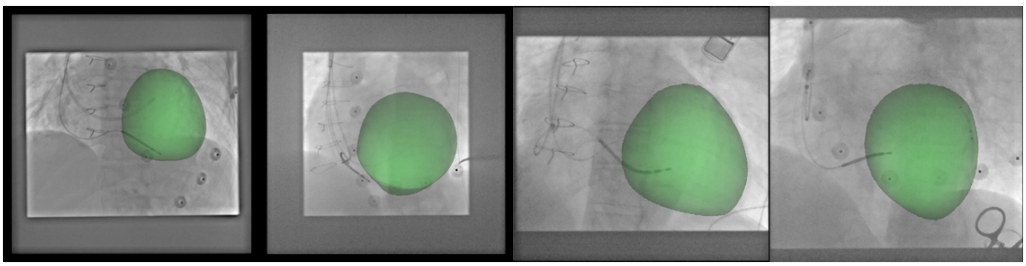


(b) DRN; W_1, D_1: Successful registrations

Figure 6.9: Effects of DR on the target domain (clinical patient X-ray data) on the RAO angulation range. Baseline qualitative results with fixed seeds (W_1 and D_1) of the (a) standard (SN) and the (b) domain randomized network (DRN).



(a) SN; W_1, D_1: Failed registrations



(b) DRN; W_1, D_1: Successful registrations

Figure 6.10: Effects of domain randomization on the target domain (clinical patient X-ray data) on the LAO angulation range. Baseline qualitative results with fixed seeds (W_1 and D_1) of the (a) standard (SN) and the (b) domain randomized network (DRN).

was trained on more synthetic data, resulting in higher overfitting to the synthetic training domain of DRRs. The DRN has, however, provided great inprovement. The robustness measure distributions are greatly compressed, the medians and means are substantially reduced, and the number of outliers has decreased significantly. The results indicate, that the DRN generalizes substantially better to unseen data, i.e., clinically acquired X-ray images. A sample set of model-X-ray overlay images of failed registrations of the SN and of the same cases succeeding with the DRN is shown in Figure 6.10.

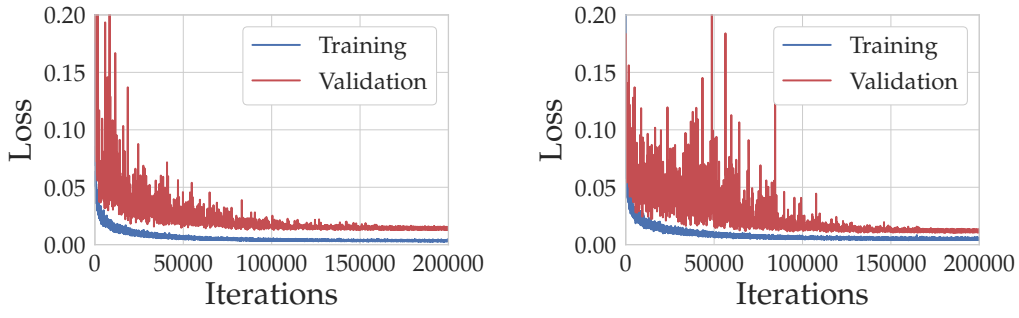
6.3.3 Parameter Variation

Two experiments were performed to assess the influence of training parameters on the evaluation results: 1) the seed for random network weight initialization was varied, with a fixed data order (W_i, D_1) and 2) the training data order seed was varied, with fixed weight initialization (W_1, D_i). Both experiments were performed for all three CRT angulation ranges, i.e., AP, RAO, and LAO, with both the SN and the DRN.

6.3.3.1 AP Range

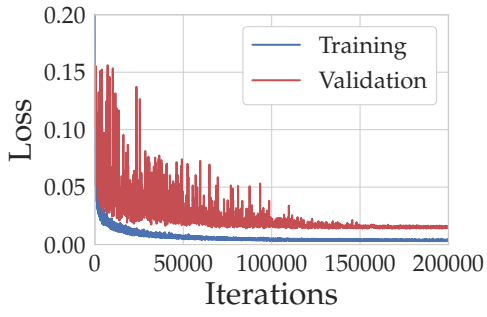
Both experiments, i.e., weight initialization seed variation and data order seed variation, had highly similar outcomes, see Figure 6.11. The training curves are highly similar among different initialization parameters. There is no substantial difference between the training curves of SNs and DRNs. The validation curves, however, show high variation. The curves converge to the same value throughout parameter combinations, but they are closer to the training curves for DRNs, suggesting less overfitting, as described in Section 6.3.2.

Results on the synthetic data show that, as expected, all training parameters with both networks (SN and DRN) result in highly similar performance (minimal improvement with the DRN), because no domain transfer is required, for a selection of results see Table 6.2 (a). For the synthetic results of all configurations see Tables A.1 and A.2.

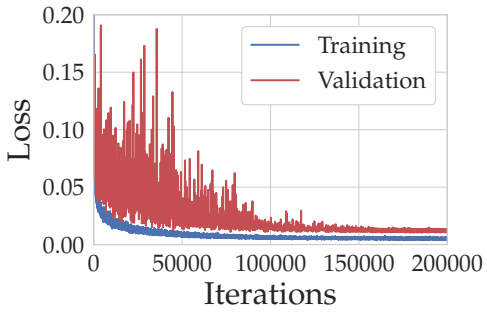


(a) SN; W_2, D_1

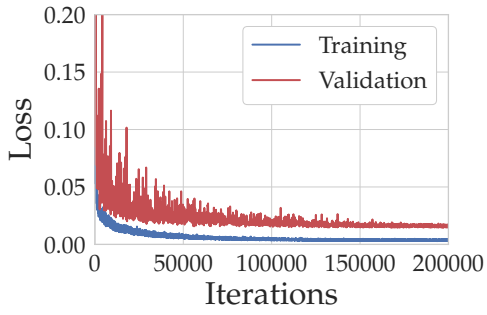
(b) DRN; W_2, D_1



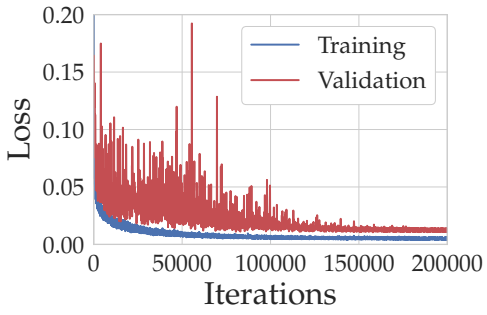
(c) SN; W_3, D_1



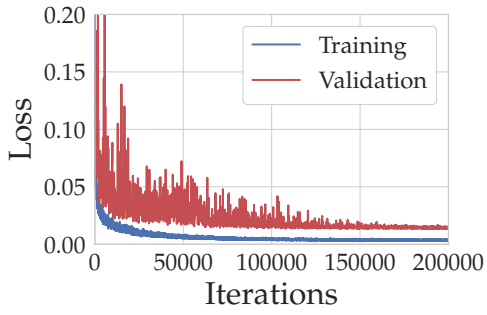
(d) DRN; W_3, D_1



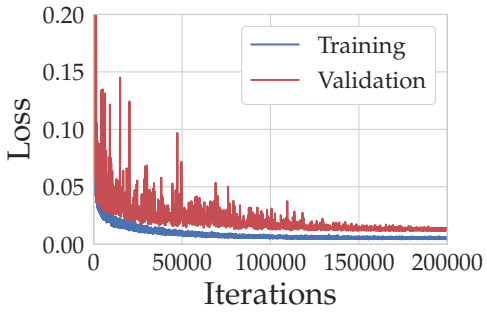
(e) SN; W_1, D_4



(f) DRN; W_1, D_4



(g) SN; W_1, D_5



(h) DRN; W_1, D_5

Figure 6.11: A selection of training and validation curves of the standard (SNs) and the domain randomized networks (DRNs), with weight initialization W_i and data order seeds D_i varied on the AP angulation range.

Table 6.2: Registration accuracy in mm for a selection of configurations with both the weight initialization W_i and the data order seeds D_i varied on synthetic DRR data of the standard (SNs) and domain randomized networks (DRNs) on the three angulation ranges AP, RAO, and LAO.

(a) AP

		SN/DRN				
	Start	W_1, D_1	W_2, D_1	W_3, D_1	W_1, D_4	W_1, D_5
Mean	22.41	2.80/2.65	2.66/2.66	2.70/2.65	2.80/2.78	2.70/2.62
StD.	10.60	2.17/2.05	2.27/2.12	2.18/2.10	2.21/2.23	2.19/2.07
Median	21.27	2.15/2.16	1.99/2.14	2.15/2.14	2.25/2.19	2.08/2.10
90 %	37.12	5.54/5.00	5.50/5.34	5.31/5.33	5.39/5.50	5.14/5.05

(b) RAO

		SN/DRN				
	Start	W_1, D_1	W_2, D_1	W_3, D_1	W_1, D_4	W_1, D_5
Mean	22.71	3.13/3.26	3.16/3.21	3.06/3.14	3.22/3.09	3.20/3.09
StD.	10.70	2.38/2.41	2.46/2.38	2.29/2.21	2.41/2.26	2.38/2.19
Median	21.36	2.45/2.60	2.52/2.61	2.44/2.56	2.61/2.51	2.49/2.56
90 %	36.86	6.39/6.77	6.43/6.38	6.22/6.14	6.20/6.07	6.60/6.07

(c) LAO

		SN/DRN				
	Start	W_1, D_1	W_2, D_1	W_3, D_1	W_1, D_4	W_1, D_5
Mean	22.49	2.44/2.42	2.41/2.39	2.47/2.37	2.54/2.43	2.49/2.45
StD.	10.43	1.69/1.75	1.63/1.79	1.81/1.68	1.70/1.72	1.82/1.70
Median	21.51	2.04/1.91	2.01/1.91	2.01/1.95	2.16/2.03	2.01/2.07
90 %	36.43	4.52/4.68	4.61/4.79	4.69/4.71	4.66/4.69	4.72/4.56

In the case of the clinically acquired X-ray data, the SN produces acceptable, but highly variable results. For a sample of the resulting robustness measure distributions for different parameter combinations in terms of their medians, quartiles and number of outliers, see Figure 6.12. The results for all combinations are illustrated in Figures A.7 and A.8.

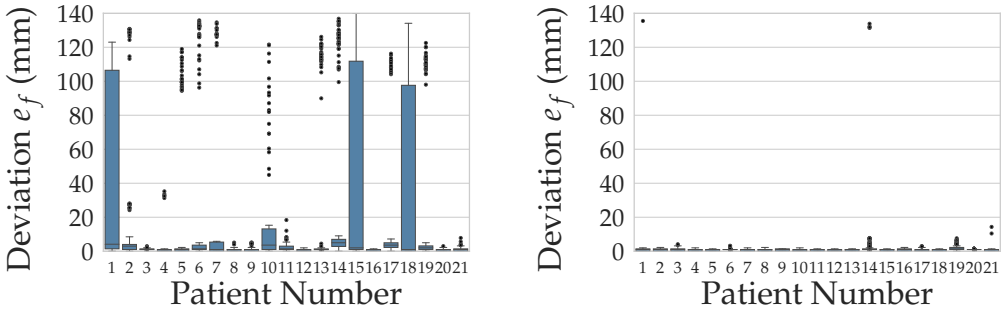
The high level of variation can be explained by the weight optimization (the training process) finishing in different local minima of the optimization space. These minima provide highly similar results on the synthetic domain, but the results can vary greatly on the target domain of clinical X-rays, as illustrated in Figure 6.1.

Focusing on individual cases, such as patients 3 and 8, good robustness and no major variation was present with the SN. The features of these images must be shared between the training and target domains, thus they can be registered without an explicit transfer. Cases such as 12 or 15 vary greatly with varying training parameters. The trained SNs are in different minima of the optimization space, and the features required for performing well on these cases are learned only in some of the setups.

The DRN produces significantly more robust and consistent results. The robustness distributions are compressed in the majority of the cases and the number of outliers is greatly reduced. The differences between the individual results are minimal. The experiments suggest that the DRN is agnostic to the variation of training parameters. The DRN trained on the AP range has shown that the transfer from synthetic to patient data is more consistent in terms of robustness than that with the SN.

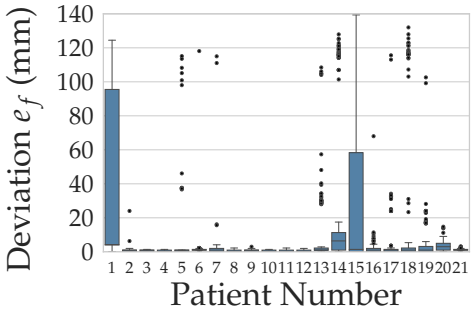
6.3.3.2 RAO Range

On the RAO range, neither the training, nor the validation curves have shown large variation for different training parameters, for a selection of training and validation curves see Figure 6.13. The complete set of results is shown in Figures A.3 and A.4. Less overfitting of the DRN can be observed, similarly to the

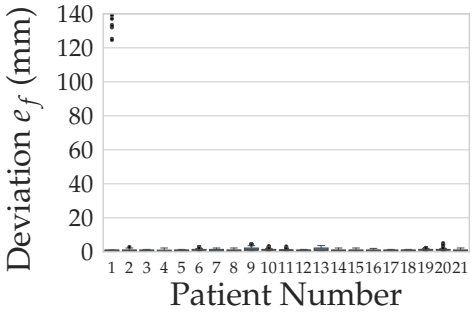


(a) SN; W_2, D_1

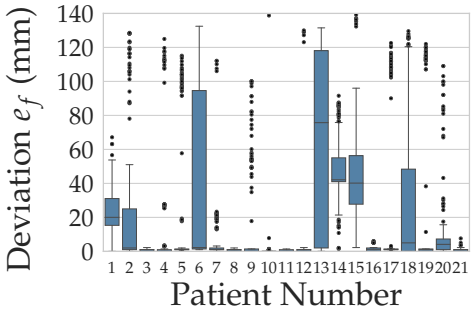
(b) DRN; W_2, D_1



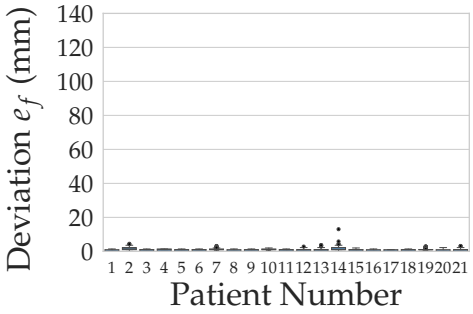
(c) SN; W_3, D_1



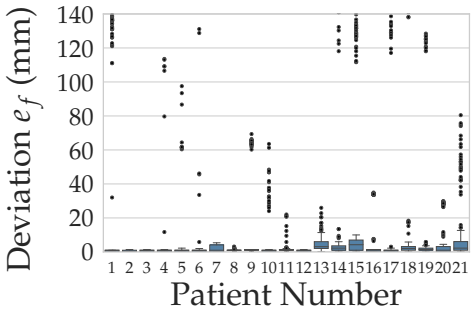
(d) DRN; W_3, D_1



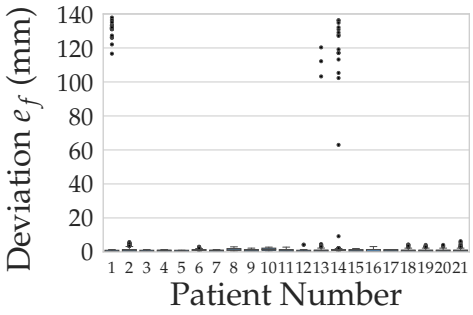
(e) SN; W_1, D_4



(f) DRN; W_1, D_4



(g) SN; W_1, D_5



(h) DRN; W_1, D_5

Figure 6.12: A selection of robustness results on clinical X-ray data of the standard (SNs) and the domain randomized networks (DRNs) with the weight initialization W_i and the training data order seeds D_i varied on the AP range.

networks trained on the AP angulation range.

Results on the synthetic data are consistent with those on the AP range, there is only a minor variation, DR does not result in improvement. For a selection of results see Table 6.2 (b). For the synthetic results of all configurations see Tables A.3 and A.4.

The clinical X-ray data experiments with the SN trained on the RAO angulation range have shown less significant variation than on the AP range, see Figure 6.14. This is consistent with the low variation of the validation curves for different training parameters. This can be explained by the networks learning similar features. Unlike in the case of the networks trained on the AP range, the optimization process (training) finishes in the same local minimum (or very similar local minima).

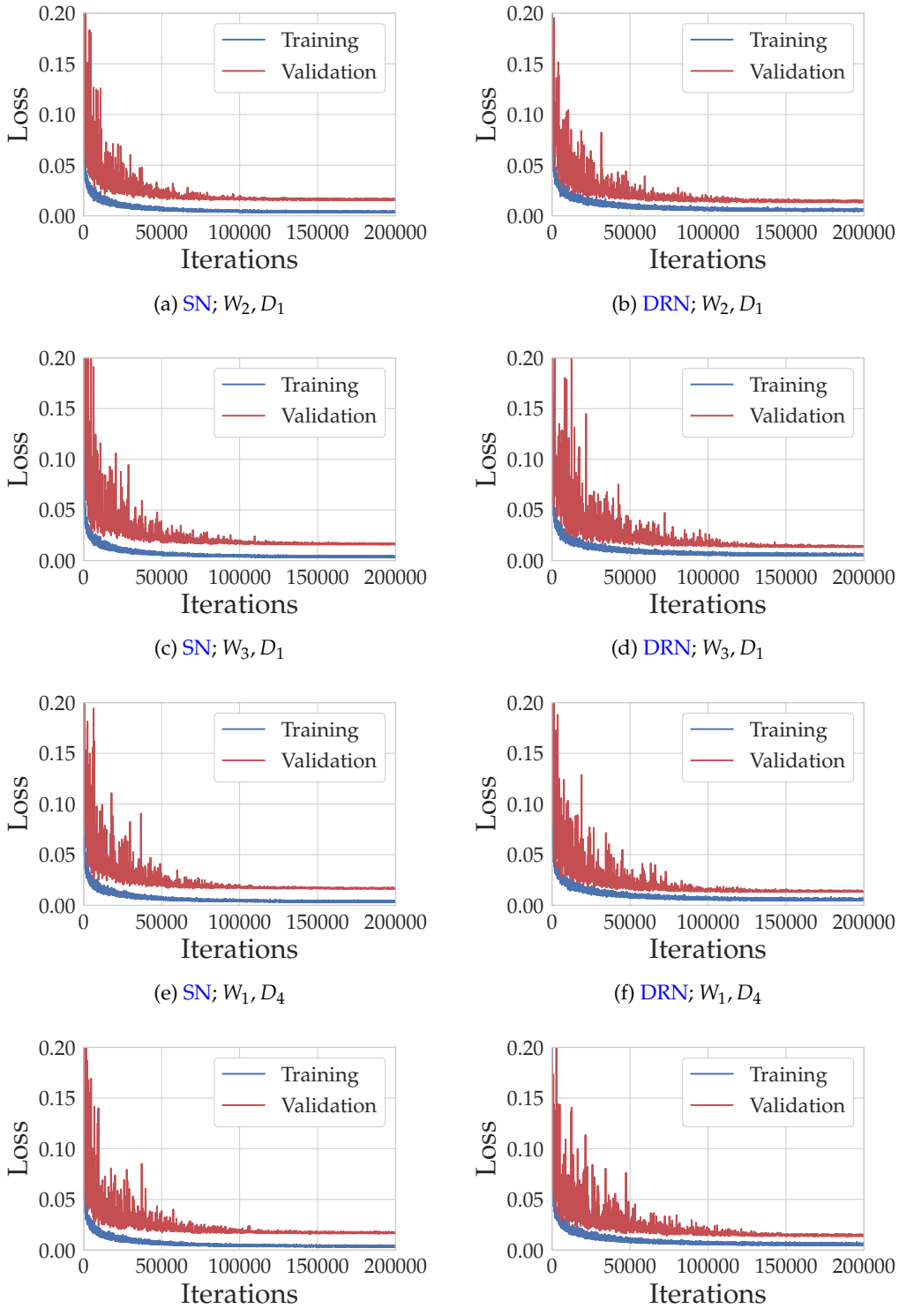
Similarly to the AP angulation range, the DRNs provide more consistent results on the clinical X-ray data. The DRNs trained on the RAO range provide very high levels of robustness. The variance of the individual patient distributions is compressed to minimal levels and the number of outliers is negligible. For the complete set of results see Figures A.9 and A.10.

6.3.3.3 LAO Range

The training and validation curves behave similarly to those of the AP angulation range, see Figure 6.15. The training curves are highly similar for all training parameter combinations. The validation curves, however, vary significantly. For the complete set of training and validation curves see Figures A.5 and A.6.

Similarly to the AP and RAO angulation ranges, the synthetic results have minimal variation. The DRNs do not improve the results either. A selection of results is displayed in Table 6.2 (c). For the complete set of synthetic results see Tables A.5 and A.6.

The clinical X-ray data results with the SNs show the lowest robustness and are the least consistent, if the training parameters change, see Figure 6.16.



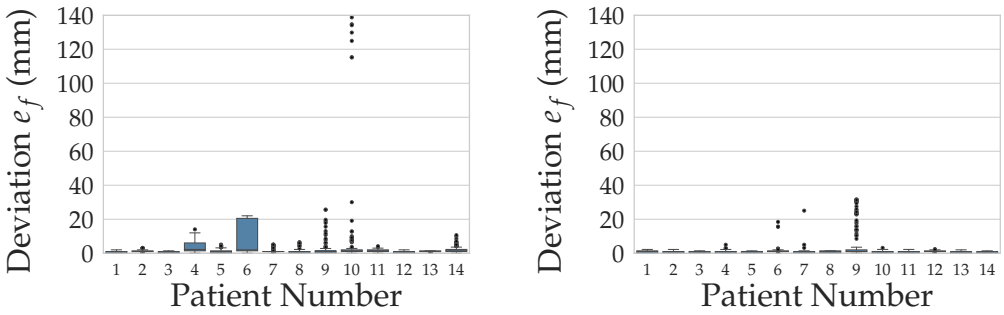
(a) SN; W_2, D_1 (b) DRN; W_2, D_1

(c) SN; W_3, D_1 (d) DRN; W_3, D_1

(e) SN; W_1, D_4 (f) DRN; W_1, D_4

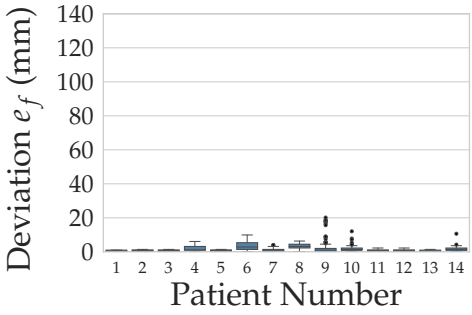
(g) SN; W_1, D_5 (h) DRN; W_1, D_5

Figure 6.13: A selection of training and validation curves of the standard (SNs) and the domain randomized networks (DRNs), with weight initialization W_i and data order seeds D_i varied on the RAO angulation range.

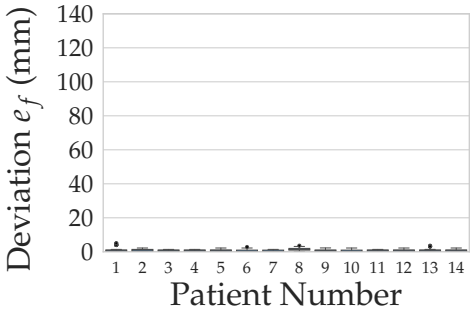


(a) SN; W_2, D_1

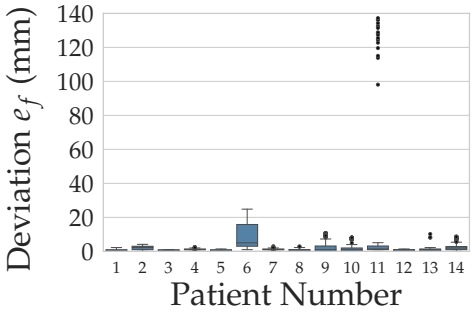
(b) DRN; W_2, D_1



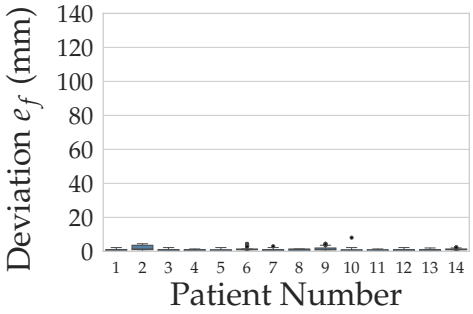
(c) SN; W_3, D_1



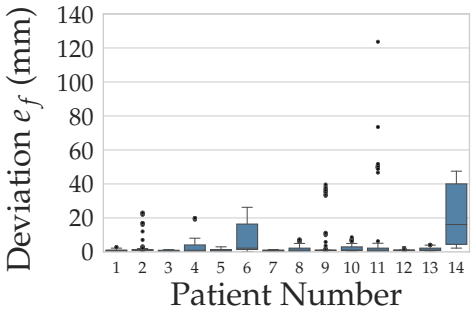
(d) DRN; W_3, D_1



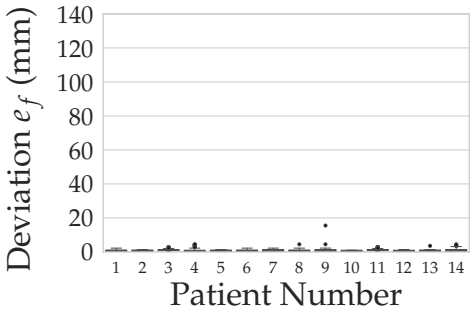
(e) SN; W_1, D_4



(f) DRN; W_1, D_4



(g) SN; W_1, D_5



(h) DRN; W_1, D_5

Figure 6.14: A selection of robustness results of the standard (SNs) and the domain randomized networks (DRNs) with the weight initialization W_i and the training data order seeds D_i varied on the RAO angulation range.

There is no case that would provide consistently good robustness throughout the parameter combinations.

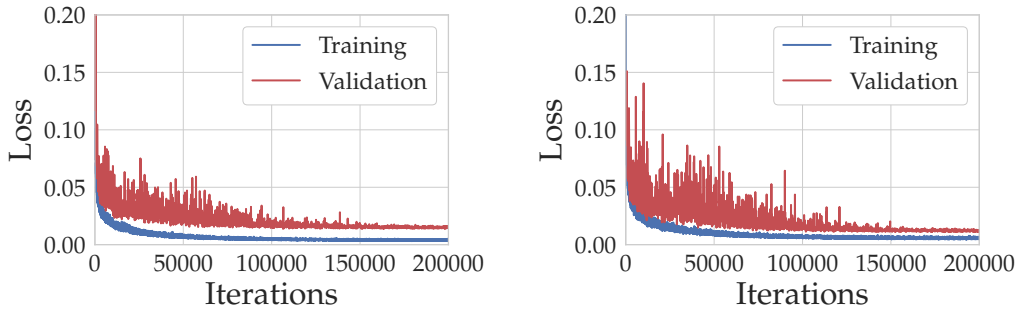
The robustness and consistency increase significantly with DR performed, see Figure 6.16. The distributions are greatly compressed for every individual case and the number of outliers is reduced substantially. The improvement is the most substantial, compared to the other two angulation ranges. However, the DRN results are still the least consistent on the LAO range. The complete set of results is displayed in Figures A.11 and A.12.

6.4 Discussion and Conclusion

This chapter has explored the robustness of transfer from the synthetic training to the clinical target domain of the imitation learning-based cardiac image registration system described in Chapter 5. Extensive experiments were conducted to evaluate the influence of training parameters, i.e., the weight initialization of the neural network representing the imitation learning agent and the training data order. It is demonstrated that DR can increase the transfer performance, providing more consistency.

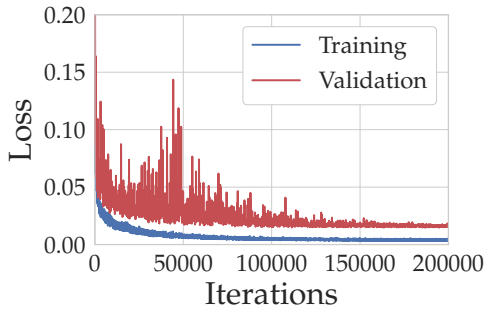
The imitation learning-based system, presented in Chapter 5, is trained entirely on artificially generated, synthetic X-ray images, DRRs. However, in this chapter, the DRR generation process was altered by performing DR, i.e., 1) varying the intensity transfer function and 2) adding artificial image borders (representing collimation) to the images. DR is an effective technique to increase the generalizability of a system trained on artificially generated data. Instead of explicitly modeling the data distribution of the target domain, e.g., with a high quality simulator (DRR renderer), a large variety of dissimilar, unrealistic images is generated. The large variation in the generated training data can enhance transfer, i.e., performance on the target data domain, because the trained network will perceive the target data as just another possible variation.

The DRN (trained on DR DRR data) and the SN (trained on standard DRR data) were trained on 1711 CT volumes in total, 1009 more CT volumes than the

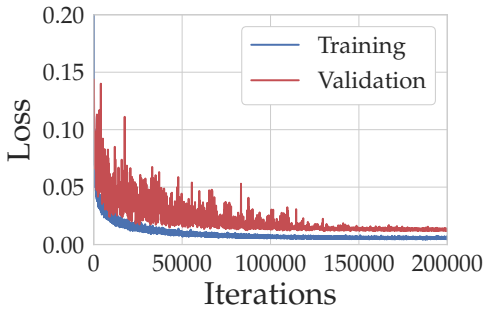


(a) SN; W_2, D_1

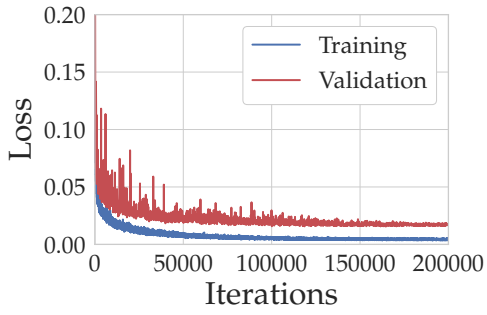
(b) DRN; W_2, D_1



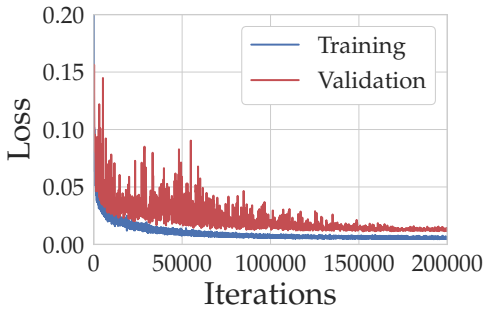
(c) SN; W_3, D_1



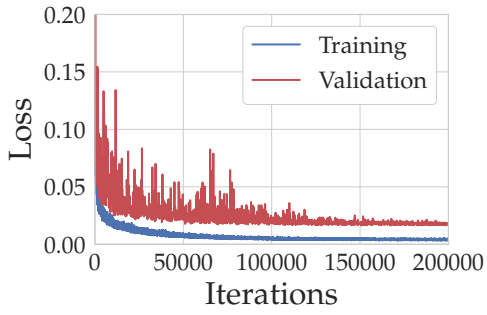
(d) DRN; W_3, D_1



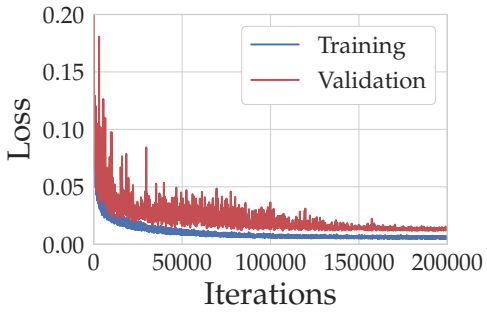
(e) SN; W_1, D_4



(f) DRN; W_1, D_4

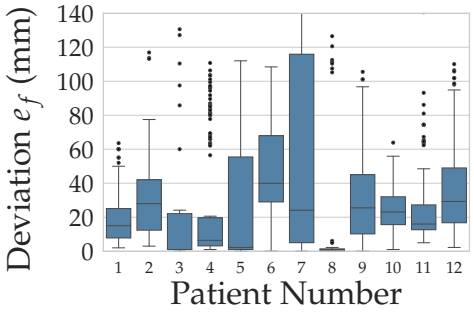


(g) SN; W_1, D_5

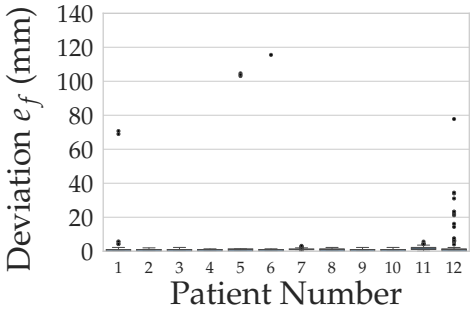


(h) DRN; W_1, D_5

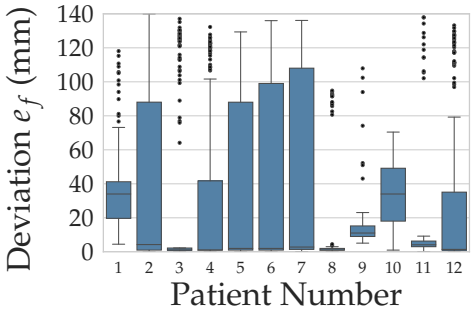
Figure 6.15: A selection of training and validation curves of the standard (SNs) and the domain randomized networks (DRNs), with weight initialization W_i and data order seeds D_i varied on the LAO angulation range.



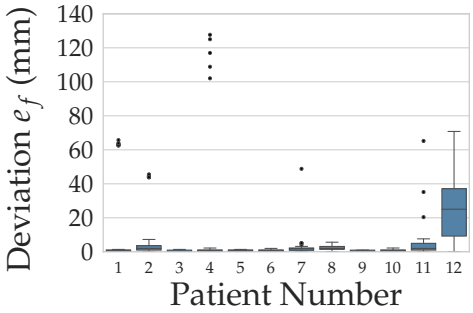
(a) SN; W_2, D_1



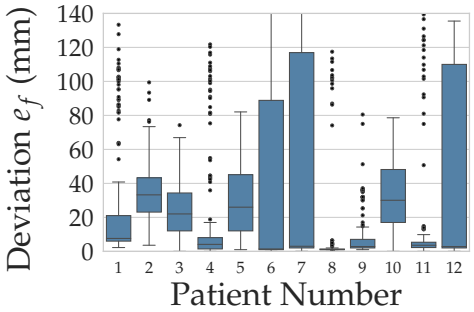
(b) DRN; W_2, D_1



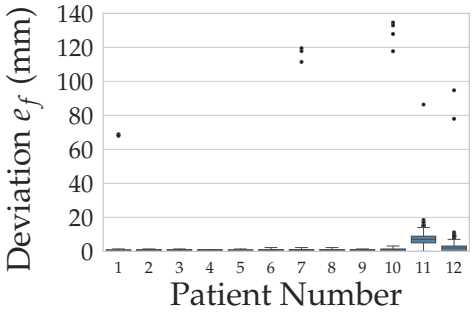
(c) SN; W_3, D_1



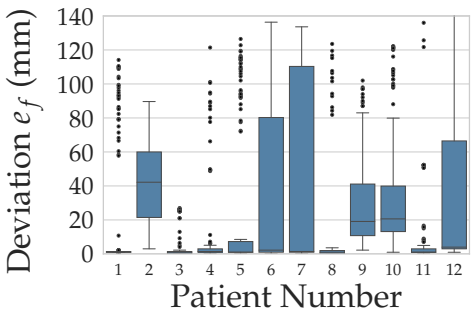
(d) DRN; W_3, D_1



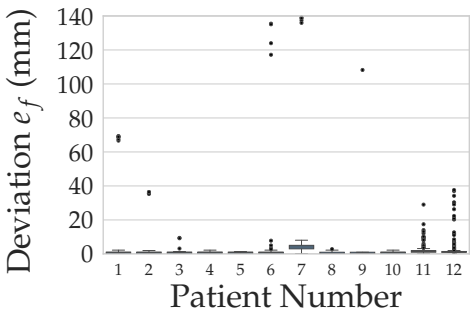
(e) SN; W_1, D_4



(f) DRN; W_1, D_4



(g) SN; W_1, D_5



(h) DRN; W_1, D_5

Figure 6.16: A selection of robustness results of the standard (SNs) and the domain randomized networks (DRNs) with the weight initialization W_i and the training data order seeds D_i varied on the LAO angulation range.

SN in Chapter 5 was trained on. The number of training iterations was doubled to 200 000 to account for the larger amount of training data and ensure convergence. As described in Chapter 5, separate networks were trained and evaluated on the three CRT angulation ranges, i.e., AP, RAO, and LAO.

Training on substantially more synthetic data, resulted in further overfitting of the SN to the synthetic domain. Registration accuracy on the synthetic data domain (DRRs) has increased marginally, however, robustness on clinically acquired X-ray data has decreased. Reduced performance of a network, if trained on more data is counter-intuitive. However, if the training and target data domains differ, the more data the network is trained on, the more it will overfit to the training domain and performance on the target domain will suffer.

It was shown in this chapter that DR can alleviate the problem of overfitting to the training domain of DRR images and provide high performance on the target domain of clinically acquired X-rays. Exeriments to evaluate robustness were performed on 21 patient datasets for the tree CRT angulation ranges. Robustness has increased in all three ranges through DR. The DRN of the RAO range has shown the highest performance and the one for the LAO range the lowest. The improvement, however, was the greatest on LAO and the lowest on RAO, suggesting that overfitting to the synthetic training data (DRRs) was the most severe on the LAO range.

A further experiment was performed to evaluate the reproducibility and consistency of the performance of the registration networks for changing training parameters. The parameters investigated were 1) the training data order and 2) the weight initialization of the neural networks. It was shown that the SNs can have highly variable results on the target domain of clinical X-ray images, however, there is minimal variation in the synthetic data (DRR) results. This can be explained, by the optimization of the weights (the training process) resulting in different local minima. The local minima provide very similar performance on the training domain, i.e., similar training losses and synthetic data performance. However, the optimization landscape on the training domain can be significantly

different. Thus, the reached minima of the training domain, might not correspond to minima of the target domain.

The DRNs have shown that the inconsistency of the target domain's results can be alleviated through DR. The DRNs provide highly similar performance in terms of robustness, the networks appear to be agnostic to the training parameters, i.e., the training data order and the weight initialization of the networks. The demonstrated robustness of the approach in this chapter can enable the interventional application of a fully automatic registration approach in cardiac procedures. In the next chapter, the conclusion of this thesis and a description of the potential extensions, the future direction of research is given.

Chapter 7

Conclusions and Future Work

The main aim of this thesis was to investigate approaches that can provide a robust interventional registration of preoperative data, i.e., magnetic resonance imaging (MRI), to interventional X-ray images for minimally invasive cardiac procedures. In the following sections, a review of the contributions of this work is given and the potential future direction of research is highlighted.

7.1 Overview of Contributions

Minimally invasive cardiac procedures, such as cardiac resynchronization therapy (CRT) delivery, can greatly benefit from the ivolvement of preoperative data, such as MRI, into the clinical workflow. The preoperative data can be utilized in preprocedural planning and for interventional image guidance. However, to enhance image guidance through the overlay of preoperative data, an accurate and robust image registration is crucial.

However, registration of these modalities is not trivial, due to the lack of cross-modality landmarks. In the majority of image guided cardiac interventions, registration between an magnetic resonance (MR) volume and an X-ray system is performed manually. Manual registration can, however, pose a significant challenge even to an experienced operator. It can be time-consuming and the results are operator-dependent. In order to provide a deterministic, robust registration, automatic approaches are much desired.

The aims are in automatic approaches for cardiac procedures, i.e., CRT, that they integrate into current clinical workflows with minimal interference, thus not relying on 1) fiducial markers, 2) interventional tools, and 3) novel MRI sequences. The following key contributions were made to advance the current state-of-the-art:

- A registration approach between reconstructed coronary veins from two X-ray angiography (XA) acquisitions and a mesh model of the left ventricle (LV), based on the adjacency constraint of the two structures given by anatomy through globally optimal iterative closest point (GO-ICP);
- The concept of superabundant reconstruction, in that all possible epipolar correspondences are reconstructed in 3D;
- An outlier tolerant adjacency-based registration approach that can register the superabundant reconstruction of the coronary veins to the LV model through GO-ICP with dynamic outlier rejection;
- The evaluation of cardiac registration through position of previously deployed implants, i.e., artificial valves;
- An imitation learning-based framework for model-to-X-ray registration, trained entirely on synthetic data;
- Analysis of transfer performance from the synthetic domain of digitally reconstructed radiographs (DRRs) to the target domain of clinical X-ray images, depending on training parameters;
- Domain randomization (DR) to enhance transfer performance, if systems trained entirely on synthetic data.

In Chapter 3, an overview of the prior work in cardiac registration, with specific focus on CRT, was provided. The challenge of no cross-modality landmarks being available between cardiac MR and X-ray images was approached with a number of methods. Besides manual registration, automatized approaches

were developed. These are either relying on external markers, tracking systems, endovascular devices, or anatomical structures. These approaches had significant interference with the standard clinical workflow, are not suitable for cine MR to X-ray registration, and a number of them has not been evaluated on in vivo data or in a clinical setting. Furthermore, the currently available clinical approach, manual registration, was evaluated to show the inter-user variability of registration results, to 1) confirm that an automatic approach is required and 2) that manual registration is not suitable to define ground truth (GT) transformations.

An approach was proposed in Chapter 4 that relies on the adjacent geometric relation of the coronary veins and the LV. Superabundant reconstruction of the coronary veins is performed from two XA acquisitions. In superabundant reconstruction, all possible epipolar correspondences are reconstructed in 3D, resulting in an outlier rich point cloud. The superabundant point cloud is registered to the LV mesh model extracted from the preoperative cine MR acquisition through the GO-ICP approach. The method was evaluated qualitatively and quantitatively on a self-created phantom dataset, a clinical computed tomography angiography (CTA) volume, and clinical CRT datasets. The method has demonstrated high accuracy on the phantom and CTA datasets. Clinical experts have rated the registration results qualitatively on the CRT data as either no corrections, or minimal corrections were necessary for interventional guidance. Furthermore, quantitative evaluation was performed on one dataset that had artificial valves implanted. The MR centerpoints of the valves were compared to the centerpoints reconstructed from the two X-ray images, showing adequate accuracy.

The adjacent anatomical model-based approach solves the registration problem between a preoperative modality and the interventional X-ray system, however, it relies on high quality XA acquisitions of the coronary veins. The main challenge is posed by the anatomy of the coronary veins. The coronary sinus (CS) can be wide, preventing total occlusion thereof, thus a sufficient contrast agent fill. Additionally, the number of available coronary veins can vary. These factors can prevent the reconstruction of a superabundant point cloud of sufficient quality

for registration. To overcome the posed challenges, an approach was proposed in Chapter 5, based on an imitation learning agent. The method relies on a single, non-contrasted X-ray acquisition. The agent is trained exclusively on synthetic data, DRRs, generated from computed tomography (CT) volumes. Quantitative evaluation has shown high accuracy and robustness on synthetic data of the same domain as the training data. High levels of robustness were demonstrated on clinical CRT cases for anterior-posterior (AP) and right anterior oblique (RAO) angulations. Acquisitions of left anterior oblique (LAO) angulations have shown substantially lower robustness, however, acceptable for the majority of cases. The lower performance on the clinical data can be explained through overfitting to the training domain of DRR data and no explicit transfer being performed to the target domain.

To enhance the performance of the imitation learning agent on the target domain of clinical X-ray acquisitions, Chapter 6 introduced the concept of DR to the context of medical imaging. DR generated DRRs with randomly varying intensity transfer functions and borders, to model intensity variations and collimation respectively. The imitation learning agent was trained on the DR DRRs. The agent without DR provided highly variable robustness, if training parameters, i.e., neural network (NN) weight initialization and training data order have changed. However, the DR results were agnostic to the change of training parametes, showing considerably lower variation in transfer performance. It was shown that the agent with DR can provide substantially higher robustness on the target domain of clinical X-ray data on all three CRT angulation ranges.

Besides the theoretical and experimental contributions, the imitation learning agent-based registration approach has been integrated into the clinical prototype for CRT guidance, described in detail in Section 2.2.4. For images of the deployed prototype see Figure 7.1. To provide a true 3D alignment, the registration in the prototype is performed with two imaging planes in parallel. The transformations of the two planes are combined after every iteration and the combined transformation is applied to the mesh model. The operating clinician can make

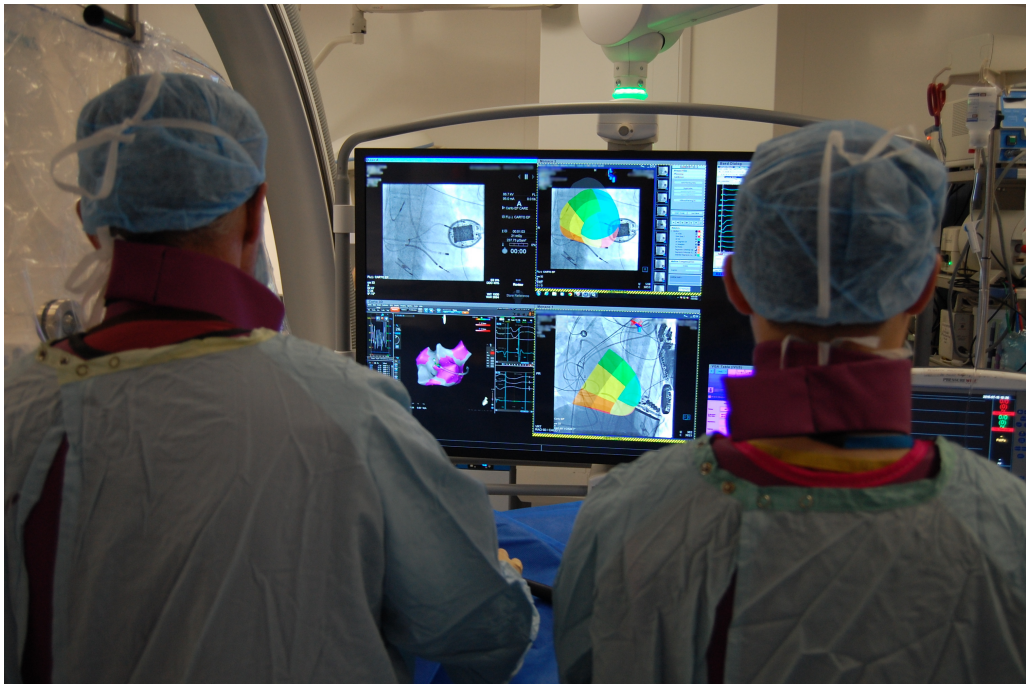
adjustments to the registration result through an interactive, manual 3D/2D registration interface. The registration framework was tested in one clinical case, achieving a successful registration with two planes (RAO-LAO), supporting the implantation of the LV lead. The registration time in the clinical setting has increased to approximately 20 s, due to performing the registration in two planes and ensuring convergence for unfavorable initializations of the two planes. The runtime is ideal for CRT interventions. No interference is caused in the standard workflow, because the registration time is significantly shorter than that of manual registration, or of the adjacent antomy-based approach.

The benefits of the prototype are planned to be evaluated in a multicenter clinical trial. The main factor investigated will be treatment outcome, however, the individual components of the prototype, amongst others the registration, can also be extensively evaluated on a large patient cohort.

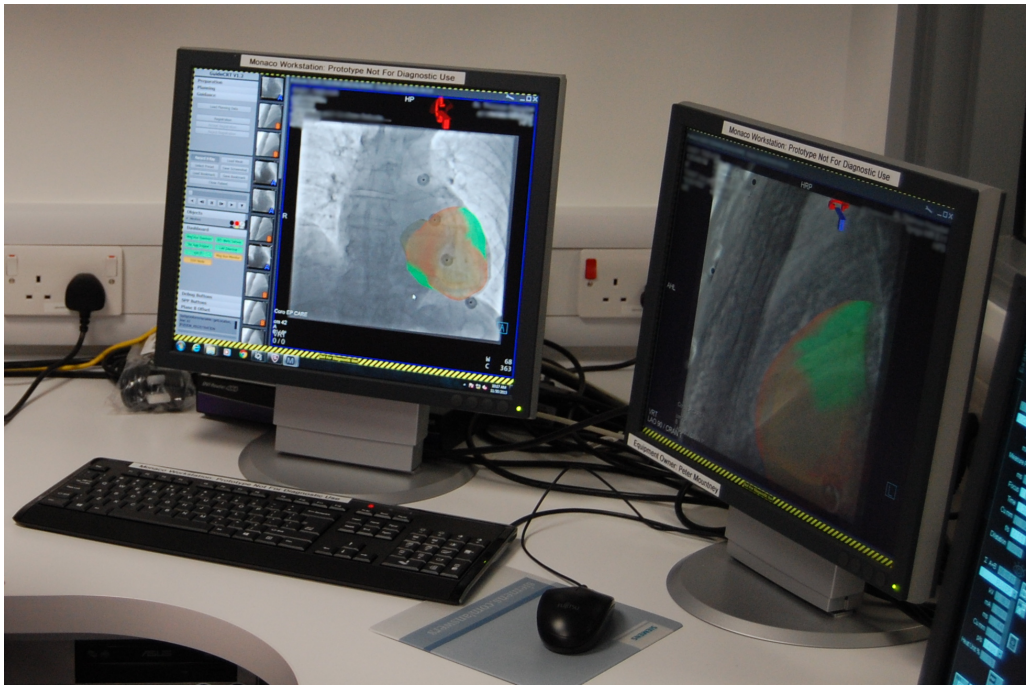
7.2 Potential Future Work

The DR imitation learning agent has demonstrated high accuracy on synthetic and high robustness on clinically acquired CRT patient data. However, quantitative measurement of accuracy on in vivo data was not possible, due to no ground truth registration being available. Accuracy could, however, be measured in animal or human cadaver experiments. Multimodal fiducial markers could be implanted onto the surface of the LV, to represent the target (CRT pacing positions). A 3D MRI volume of the animal or cadaver could be acquired. The LV model and the fiducial markers' positions can be extracted from the MRI volume and reconstructed from two X-ray acquisitions. These experiments would, however, either not include patient motion in the case of the cadaver, or use a non-human model with different anatomy.

A more clinically related accuracy evaluation method of the registration approach between cine MR and XA data would be through a CT volume acquired with contrast injection. The mesh model of the LV can be extracted from both the MR and the CT data. Additionally, the coronary veins can be segmented from



(a) Overlay in the catheterization laboratory



(b) Clinical prototype in the control room of the catheterization laboratory

Figure 7.1: The deployed clinical prototype for guiding CRT procedures (Siemens Healthcare GmbH, Forchheim, Germany). (a) Overlay images provided by the prototype for CRT guidance on the catheterization laboratory screen actively used in image guidance. (b) The control interface of the clinical prototype in the control room of the catheterization laboratory.

CT too. A registration between the two LV models can be established through, e.g., iterative closest point (ICP). The coronary veins from the CT volume can be registered to two contrasted X-ray images through a reconstruction of the coronary veins. This multistep registration can establish the transformation between cine MRI and XA. The found transformation could be used to evaluate against, however, errors of the individual transformations can accumulate and reduce the quality of the ground truth registration.

Besides providing guidance for CRT procedures, through preoperative MR data, CT image guidance is also supported. CT image guided procedures are aimed at patients with non de novo implants, that have a conventional two-lead pacemaker implanted. The implanted devices might be MR compatible, however, the MR image quality can suffer greatly from metal artifacts. The artifacts can prevent the segmentation of the LV, scar tissue identification, and quantification. However, CT volumes can be directly registered to X-rays, e.g., through bony structures, the model-to-image imitation learning agent can support CT guided procedures as well. Registering the model to X-rays has the benefit, that the target anatomy is being registered. Thus, the changing geometric relation of the heart and bony structures will not impact registration quality. However, the prior implant, the pacemaker and the two leads, can pose a challenge to the imitation learning agent-based registration. The agent was not trained explicitly with images representing such structures. The impact of implants on performance needs to be evaluated. Performance reduction is expected, that can be alleviated through, e.g., simple augmentation or performing DR of devices in the training data.

A further possible clinical application is wireless, endocardial CRT device implantation. The procedures are CT guided, however, the registration is more challenging, due to the co-implants being in the field of view (FOV), i.e., the battery and the ultrasound (US) transmitter. The battery appears mainly in the FOV of LAO and the US transmitter in the FOV of RAO acquisitions. It is to be shown whether the registration approach is robust against the devices

or further DR is required. Models of the devices can be acquired through the segmentation of a 3D modality, such as CT or cone beam computed tomography (CBCT). The models can be projected into the planes of the training images of the respective angulation ranges, with randomized geometric perturbations and intensity transformations. A system trained on such DR data can be robust against devices in the FOV, thus can be applied in wireless, endocardial pacing procedures.

The approach could find further use in other electrophysiology (EP) procedures, such as ablation for the treatment of atrial fibrillation (AFib). To enable a precise targeting of areas to ablate based on preoperative data, a robust registration approach is required. The described DR imitation learning agent is able to provide the required transformation, however, the acquisition of images that the CRT agent was trained on would significantly disrupt the standard clinical workflow. In AFib ablation, the acquired X-rays are highly concentrated onto the left atrium (LA), the LV is not, or only partially, visible. The approach would need to be retrained with the model of the LA and DRRs focusing on the LA. It is to be investigated, whether it is possible to train an agent to register based on the LA, due to its appearence in DRRs and X-rays being less prominent than that of the LV. Furthermore, the X-ray acquisitions can be of small FOVs, thus, it is to be seen, whether images of high crop can be registered reliably.

Besides EP applications, treatment of structural defects of the heart could also benefit from the DR imitation learning agent. Interventions, such as transcatheter aortic valve implantation, can be provided with a registration through the agent trained for CRT. However, for other structural diseases, such as congenital heart defects, there are further challenges to be solved. In congenital cases, the heart defects can be that severe, that the structure of interest, e.g., the LV or the LA, can vary in shape significantly. The registration agent needs to be retrained for such cases, by randomizing the shape of the structure of interest as well. This can be achieved through nonrigid transformations applied to the CT volumes used for training before segmentation, or by transforming the extracted model in parallel

to the volume after segmentation.

A further promising application of the imitation learning agent is motion compensation. After performing an initial registration, the agent could perform an iterative adjustment of the aligment in the respective imaging plane for consecutive fluoroscopy frames. The adjustment can be performed in the time frame of two consecutive fluoroscopy frames, however, this means that the adjusted position is not available in the current frame, only in the following one. The iterative adjustment can be combined with a pedicition framework, such as a Kalman [156] or a particle filter [157], to provide the registration for the consecutive frames, thus to provide a real time motion compensation. The iterative re-adjustment of the overlay through in-plane registration could also be beneficial to compensate for patient motion in the working angulation.

This thesis has presented approaches for registering preoperative image data, through models and machine learning, to interventional X-ray images of cardiac interventions. A robust registration is crucial for interventional guidance of cardiac procedures. The developed approaches, especially the DR imitation learning agent, can perform the registration with a high level of robustness, thus, having the potential to increase success rates of CRT and similar cardiac procedures.

Appendix A

Domain Randomization Results

A.1 Synthetic Data – Quantitative Results

Table A.1: Registration accuracy in mm for the weight initialization seed varied W_i, with a fixed data order seed D_1 on synthetic digitally reconstructed radiograph (DRR) data of the standard (SNs) and domain randomized networks (DRNs) on the anterior-posterior (AP) angulation range.

(a) Standard networks

	Start	W_1, D_1	W_2, D_1	W_3, D_1	W_4, D_1	W_5, D_1
Mean	22.41	2.80	2.66	2.70	2.63	2.77
StD.	10.60	2.17	2.27	2.18	2.23	2.24
Median	21.27	2.15	1.99	2.15	1.96	2.23
90 %	37.12	5.54	5.50	5.31	5.26	5.52

(b) Domain randomized networks

	Start	W_1, D_1	W_2, D_1	W_3, D_1	W_4, D_1	W_5, D_1
Mean	22.41	2.65	2.66	2.65	2.64	2.68
StD.	10.60	2.05	1.98	2.05	2.23	2.07
Median	21.27	2.16	2.14	2.14	2.08	2.14
90 %	37.12	5.00	5.34	5.33	5.07	5.24

Table A.2: Registration accuracy in mm for the data order seed varied D_i, with a fixed weight initialization seed W_1 on synthetic DRR data of the standard (SNs) and domain randomized networks (DRNs) on the AP angulation range.

(a) Standard networks

	Start	W_1, D_1	W_1, D_2	W_1, D_3	W_1, D_4	W_1, D_5
Mean	22.41	2.80	2.78	2.64	2.80	2.70
StD.	10.60	2.17	2.19	2.17	2.21	2.19
Median	21.27	2.15	2.26	2.02	2.25	2.08
90 %	37.12	5.54	5.31	5.21	5.39	5.14

(b) Domain randomized networks

	Start	W_1, D_1	W_1, D_2	W_1, D_3	W_1, D_4	W_1, D_5
Mean	22.41	2.65	2.58	2.59	2.78	2.62
StD.	10.60	2.05	1.98	2.05	2.23	2.07
Median	21.27	2.16	2.03	2.06	2.19	2.10
90 %	37.12	5.00	5.11	5.23	5.50	5.05

Table A.3: Registration accuracy in mm for the weight initialization seed varied W_i, with a fixed data order seed D_1 on synthetic DRR data of the standard (SNs) and domain randomized networks (DRNs) on the right anterior oblique (RAO) angulation range.

(a) Standard networks

	Start	W_1, D_1	W_2, D_1	W_3, D_1	W_4, D_1	W_5, D_1
Mean	22.71	3.13	3.16	3.06	3.16	3.13
StD.	10.70	2.38	2.46	2.29	2.33	2.35
Median	21.36	2.45	2.52	2.44	2.61	2.59
90 %	36.86	6.39	6.43	6.22	6.30	6.04

(b) Domain randomized networks

	Start	W_1, D_1	W_2, D_1	W_3, D_1	W_4, D_1	W_5, D_1
Mean	22.71	3.26	3.21	3.14	3.07	3.16
StD.	10.70	2.41	2.38	2.21	2.32	2.30
Median	21.36	2.60	2.61	2.56	2.45	2.54
90 %	36.86	6.77	6.38	6.14	6.13	6.29

Table A.4: Registration accuracy in mm for the data order seed varied D_i, with a fixed weight initialization seed W_1 on synthetic DRR data of the standard (SNs) and domain randomized networks (DRNs) on the RAO angulation range.

(a) Standard networks

	Start	W_1, D_1	W_1, D_2	W_1, D_3	W_1, D_4	W_1, D_5
Mean	22.71	3.13	3.08	3.17	3.22	3.20
StD.	10.70	2.38	2.28	2.60	2.41	2.38
Median	21.36	2.45	2.48	2.50	2.61	2.49
90 %	36.86	6.39	6.11	6.31	6.20	6.60

(b) Domain randomized networks

	Start	W_1, D_1	W_1, D_2	W_1, D_3	W_1, D_4	W_1, D_5
Mean	22.71	3.26	3.15	3.06	3.09	3.09
StD.	10.70	2.41	2.38	2.09	2.26	2.19
Median	21.36	2.60	2.51	2.54	2.51	2.56
90 %	36.86	6.77	6.14	5.84	6.07	6.07

Table A.5: Registration accuracy in mm for the weight initialization seed varied W_i, with a fixed data order seed D_1 on synthetic DRR data of the standard (SNs) and domain randomized networks (DRNs) on the left anterior oblique (LAO) angulation range.

(a) Standard networks

	Start	W_1, D_1	W_2, D_1	W_3, D_1	W_4, D_1	W_5, D_1
Mean	22.49	2.44	2.41	2.47	2.50	2.48
StD.	10.43	1.69	1.63	1.81	1.81	1.76
Median	21.51	2.04	2.01	2.01	2.09	2.02
90 %	36.43	4.52	4.61	4.69	4.61	4.70

(b) Domain randomized networks

	Start	W_1, D_1	W_2, D_1	W_3, D_1	W_4, D_1	W_5, D_1
Mean	22.49	2.42	2.39	2.37	2.41	2.38
StD.	10.43	1.75	1.79	1.68	1.72	1.76
Median	21.51	1.91	1.91	1.95	1.99	1.96
90 %	36.43	4.68	4.79	4.71	4.65	4.58

Table A.6: Registration accuracy in mm for the data order seed varied D_i, with a fixed weight initialization seed W_1 on synthetic DRR data of the standard (SNs) and domain randomized networks (DRNs) on the LAO angulation range.

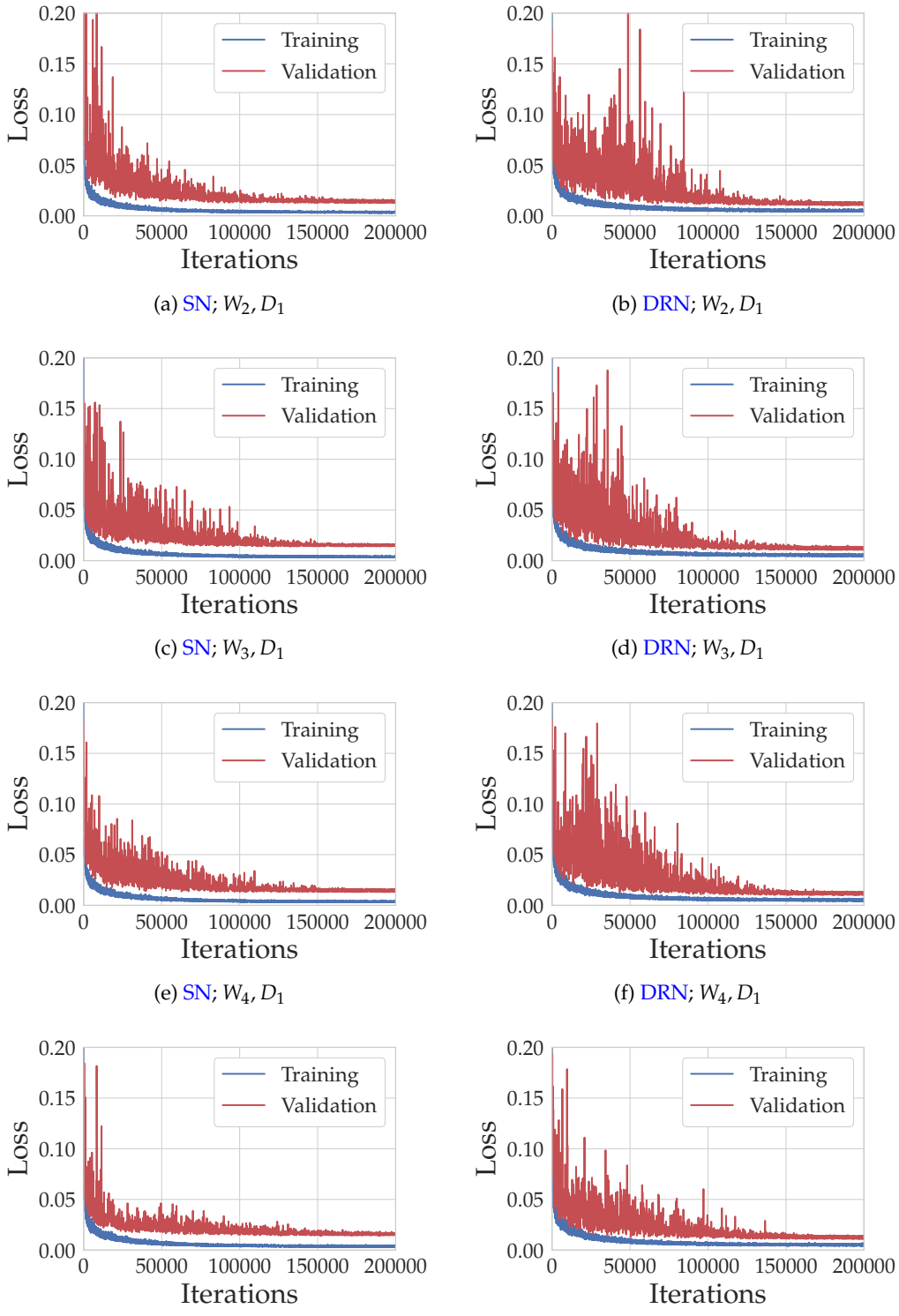
(a) Standard networks

	Start	W_1, D_1	W_1, D_2	W_1, D_3	W_1, D_4	W_1, D_5
Mean	22.49	2.44	2.57	2.58	2.54	2.49
StD.	10.43	1.69	1.84	1.78	1.70	1.82
Median	21.51	2.04	2.13	2.14	2.16	2.01
90 %	36.43	4.52	4.85	4.97	4.66	4.72

(b) Domain randomized networks

	Start	W_1, D_1	W_1, D_2	W_1, D_3	W_1, D_4	W_1, D_5
Mean	22.49	2.42	2.46	2.41	2.43	2.45
StD.	10.43	1.75	1.67	1.68	1.72	1.70
Median	21.51	1.91	2.12	1.97	2.03	2.07
90 %	36.43	4.68	4.51	4.67	4.69	4.56

A.2 Training Curves



(a) SN; W_2, D_1

(b) DRN; W_2, D_1

(c) SN; W_3, D_1

(d) DRN; W_3, D_1

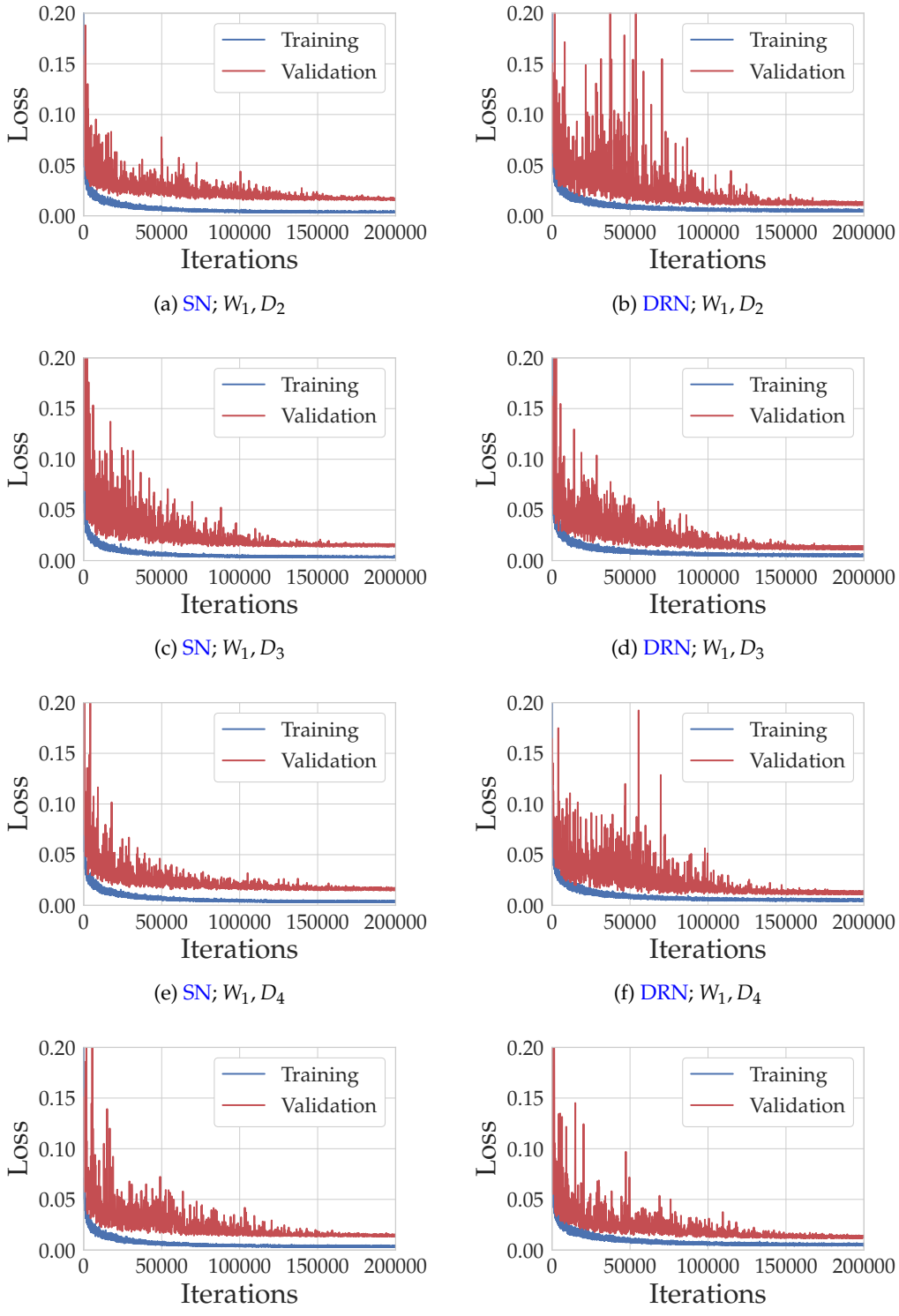
(e) SN; W_4, D_1

(f) DRN; W_4, D_1

(g) SN; W_5, D_1

(h) DRN; W_5, D_1

Figure A.1: Training and validation curves of the standard (SNs) and domain randomized networks (DRNs), with weight initialization varied W_i and a fixed data order seed D_1 on the AP angulation range.



(a) SN; W_1, D_2

(b) DRN; W_1, D_2

(c) SN; W_1, D_3

(d) DRN; W_1, D_3

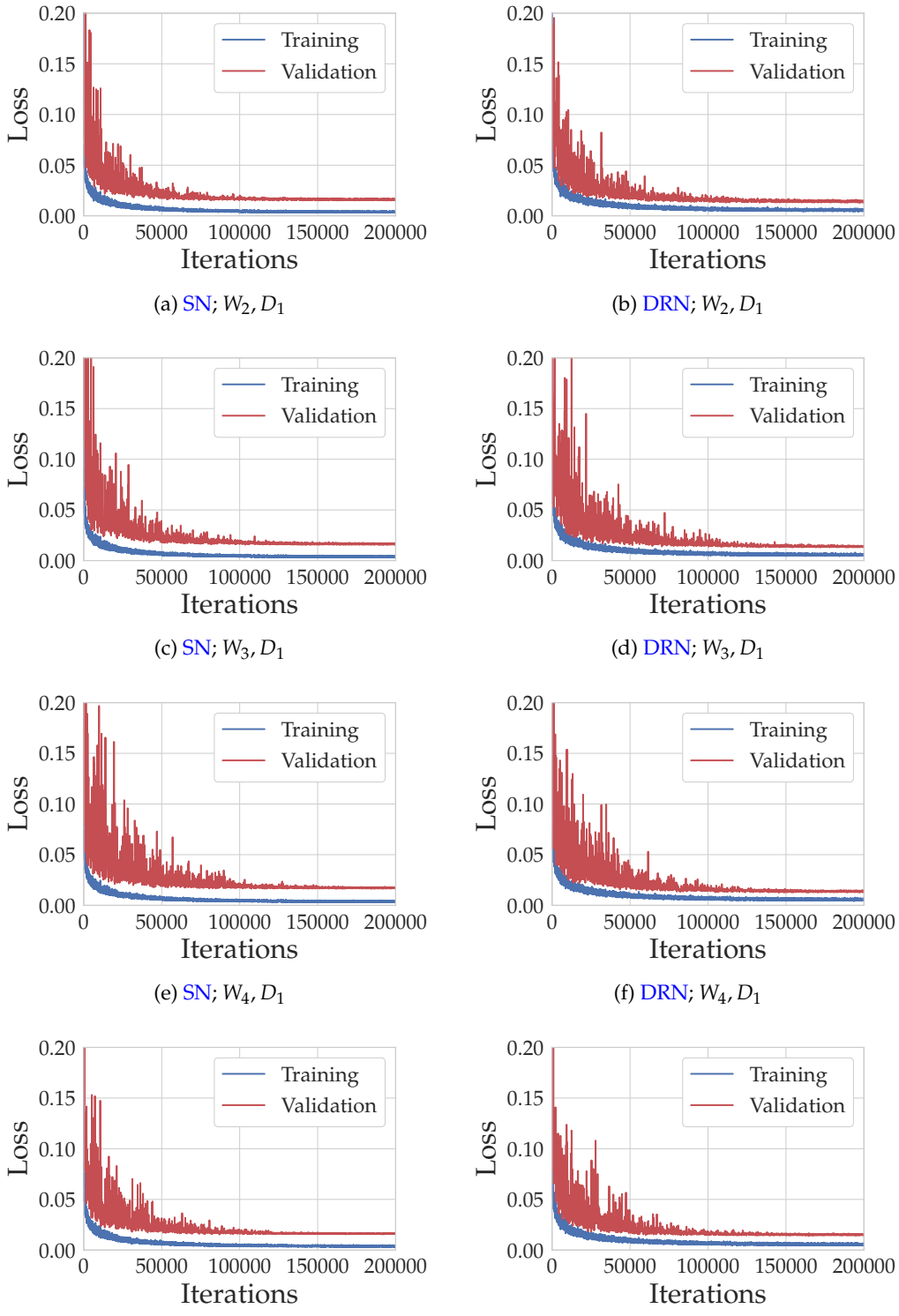
(e) SN; W_1, D_4

(f) DRN; W_1, D_4

(g) SN; W_1, D_5

(h) DRN; W_1, D_5

Figure A.2: Training and validation curves of the standard (SNs) and domain randomized networks (DRNs), with data order varied D_i and a fixed weight initialization seed W_1 on the AP angulation range.



(a) SN; W_2, D_1

(b) DRN; W_2, D_1

(c) SN; W_3, D_1

(d) DRN; W_3, D_1

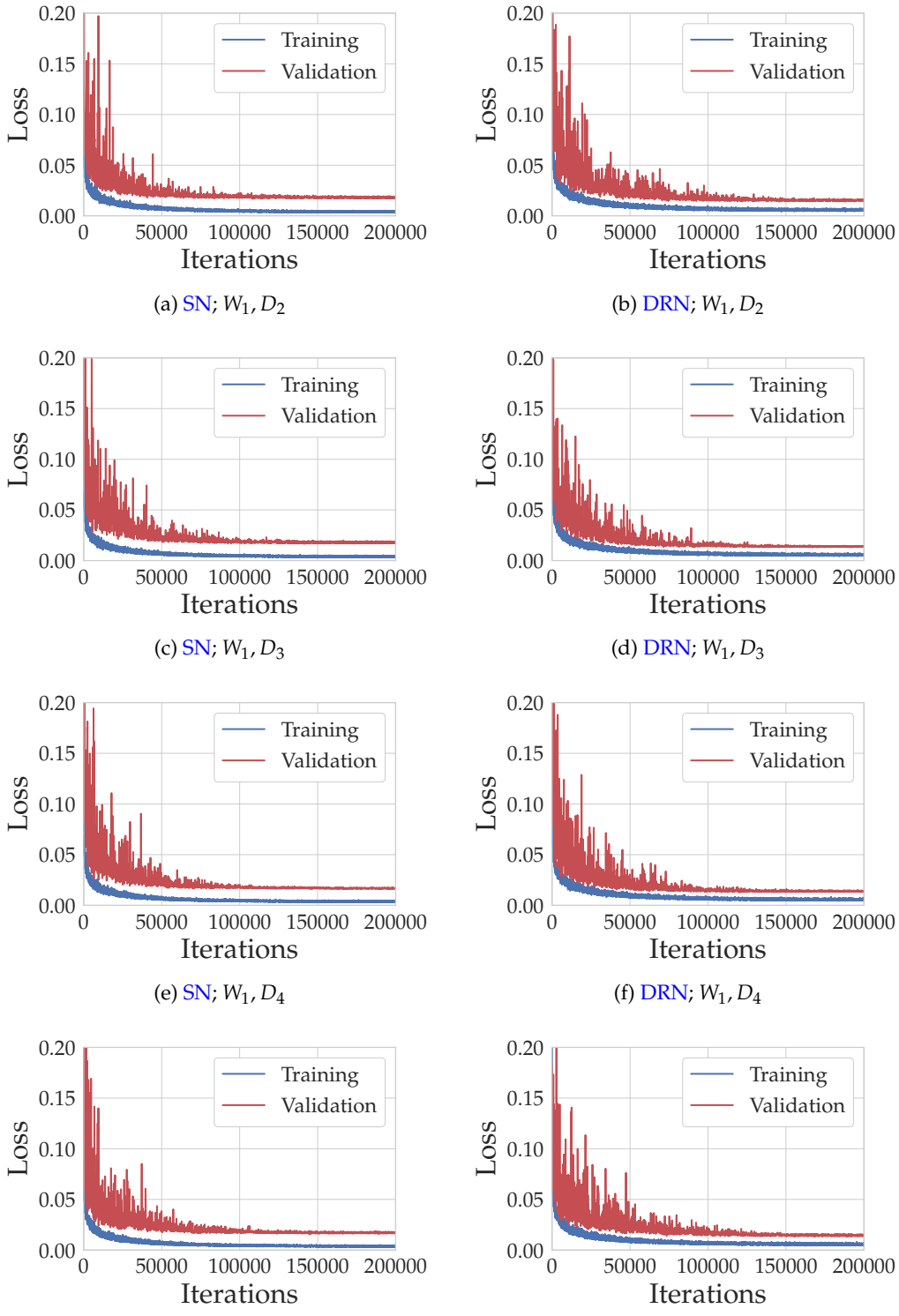
(e) SN; W_4, D_1

(f) DRN; W_4, D_1

(g) SN; W_5, D_1

(h) DRN; W_5, D_1

Figure A.3: Training and validation curves of the standard (SNs) and domain randomized networks (DRNs), with weight initialization varied W_i and a fixed data order seed D_1 on the RAO angulation range.



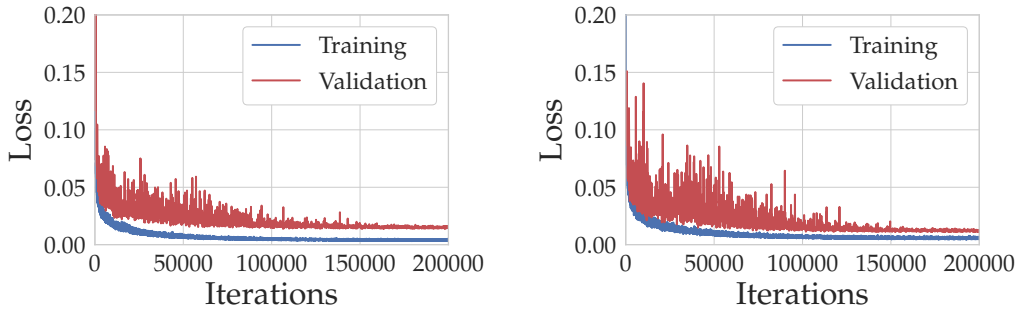
(a) SN; W_1, D_2 (b) DRN; W_1, D_2

(c) SN; W_1, D_3 (d) DRN; W_1, D_3

(e) SN; W_1, D_4 (f) DRN; W_1, D_4

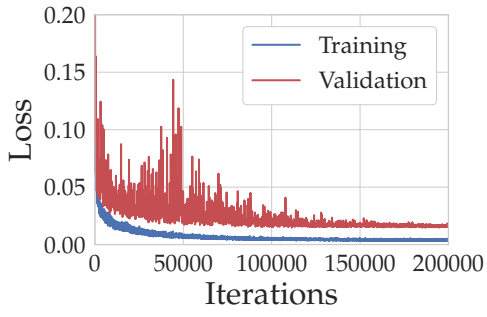
(g) SN; W_1, D_5 (h) DRN; W_1, D_5

Figure A.4: Training and validation curves of the standard (SNs) and domain randomized networks (DRNs), with data order varied D_i and a fixed weight initialization seed W_1 on the RAO angulation range.

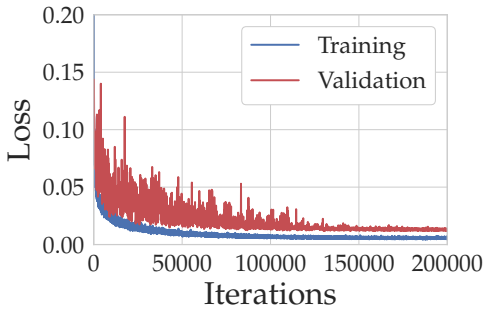


(a) SN; W_2, D_1

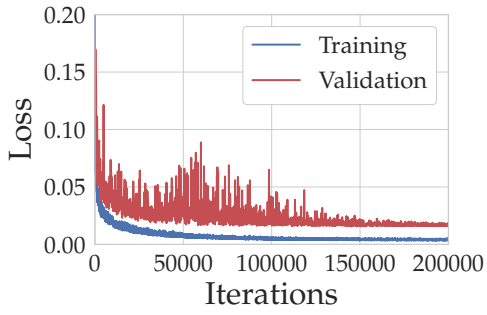
(b) DRN; W_2, D_1



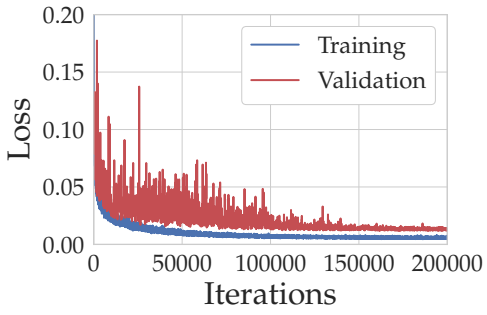
(c) SN; W_3, D_1



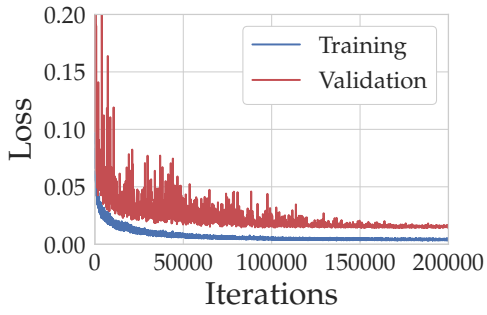
(d) DRN; W_3, D_1



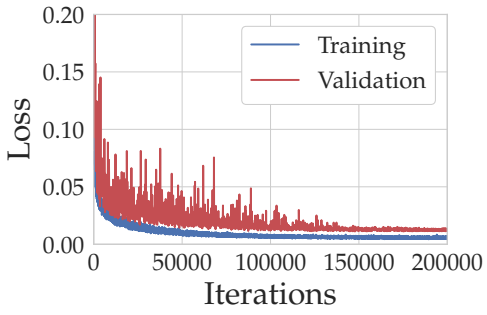
(e) SN; W_4, D_1



(f) DRN; W_4, D_1

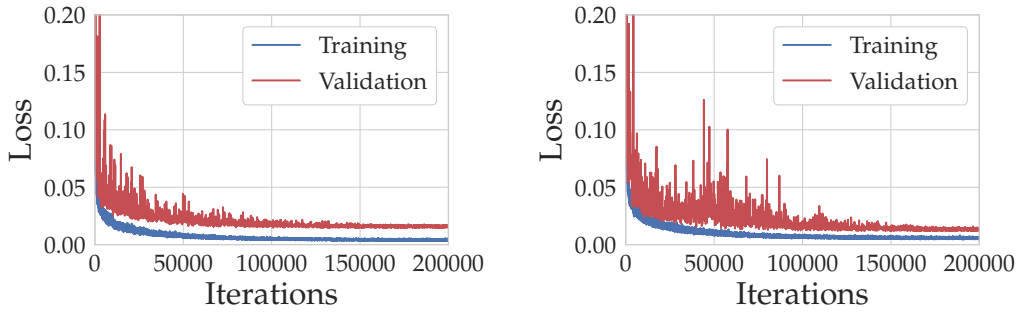


(g) SN; W_5, D_1



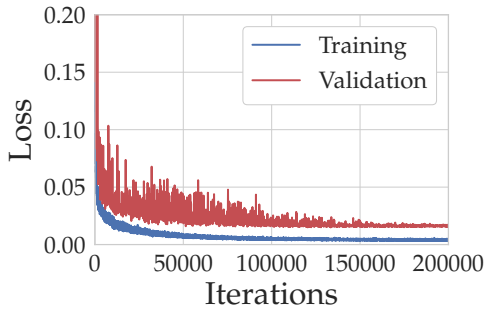
(h) DRN; W_5, D_1

Figure A.5: Training and validation curves of the standard (SNs) and domain randomized networks (DRNs), with weight initialization varied W_i and a fixed data order seed D_1 on the LAO angulation range.

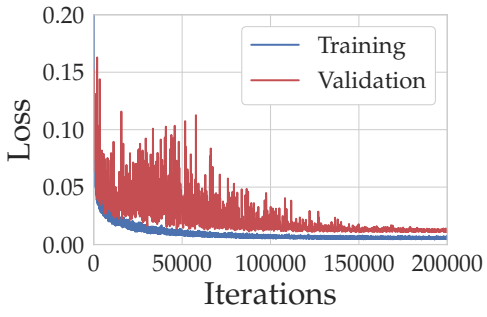


(a) SN; W_1, D_2

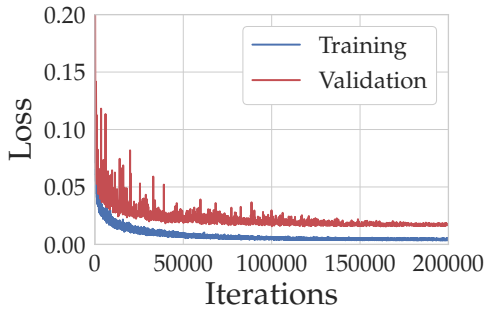
(b) DRN; W_1, D_2



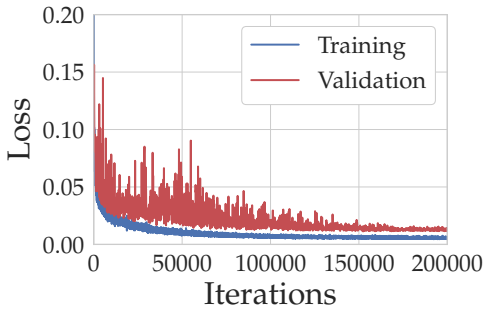
(c) SN; W_1, D_3



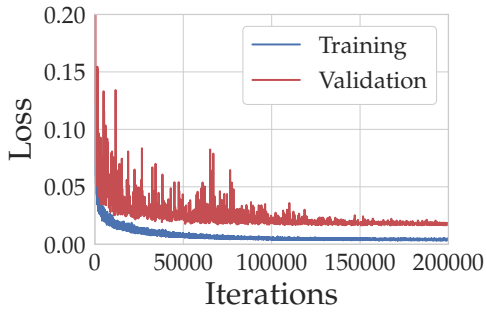
(d) DRN; W_1, D_3



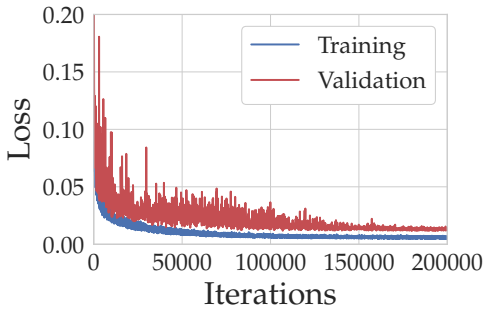
(e) SN; W_1, D_4



(f) DRN; W_1, D_4



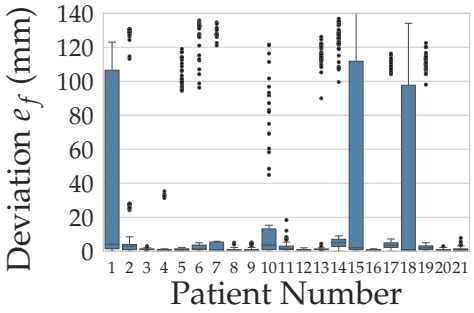
(g) SN; W_1, D_5



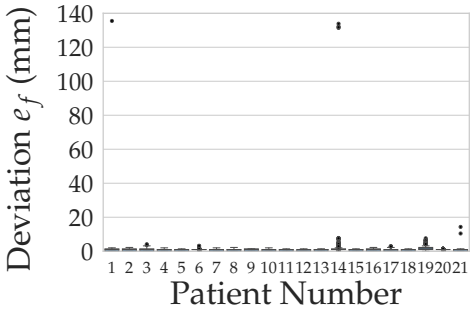
(h) DRN; W_1, D_5

Figure A.6: Training and validation curves of the standard (SNs) and domain randomized networks (DRNs), with data order varied D_i and a fixed weight initialization seed W_1 on the LAO angulation range.

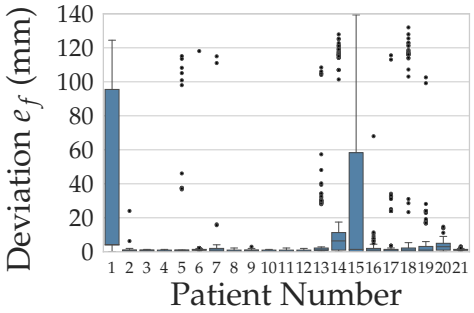
A.3 Robustness Distributions



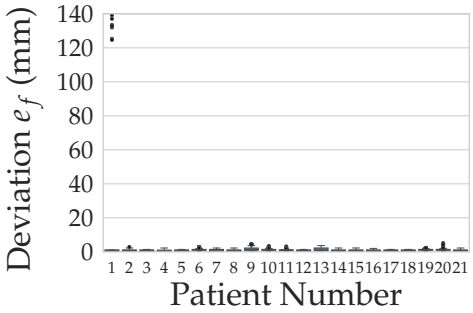
(a) SN; W_2, D_1



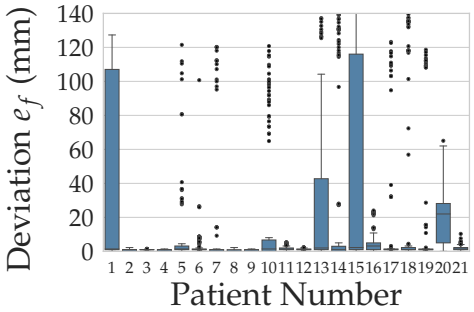
(b) DRN; W_2, D_1



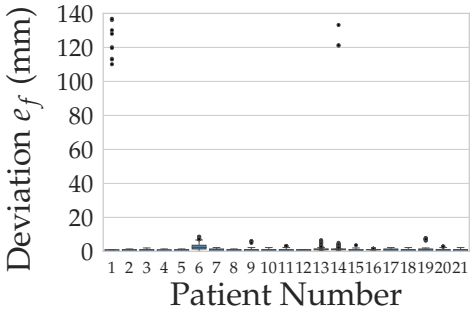
(c) SN; W_3, D_1



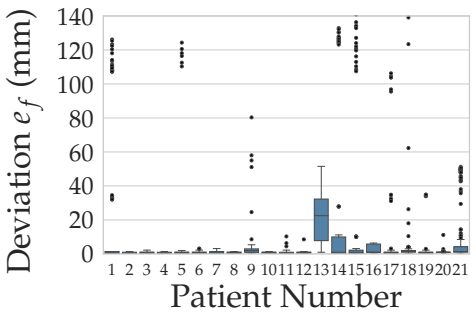
(d) DRN; W_3, D_1



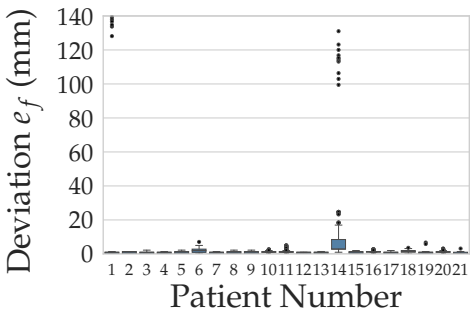
(e) SN; W_4, D_1



(f) DRN; W_4, D_1

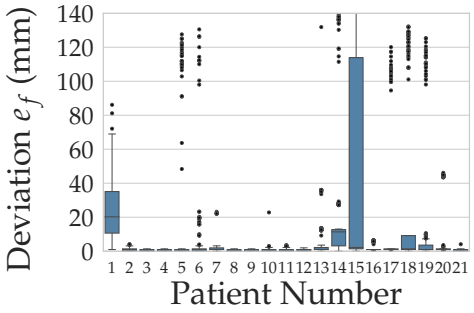


(g) SN; W_5, D_1

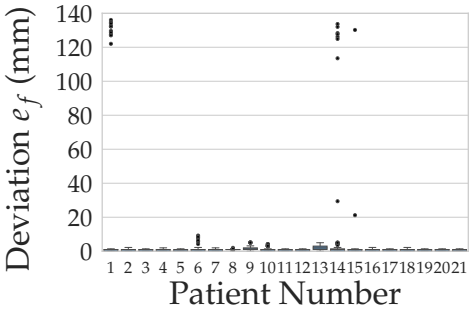


(h) DRN; W_5, D_1

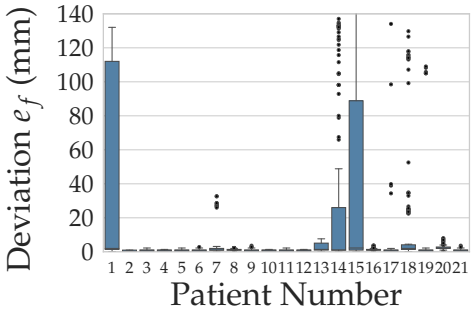
Figure A.7: Robustness results of the weight initialization seed variation W_i with a fixed training data order seed D_1 of the standard (SNs) and the domain randomized networks (DRNs) on the AP angulation range.



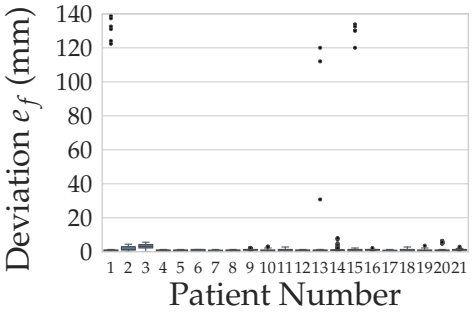
(a) SN; W_1, D_2



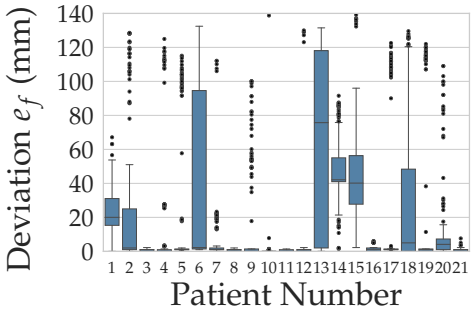
(b) DRN; W_1, D_2



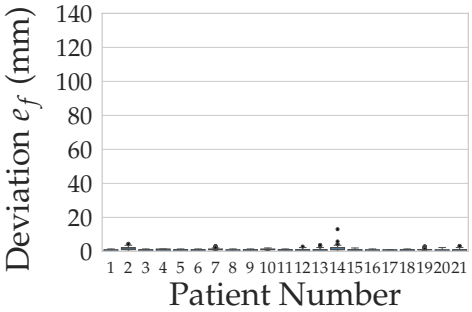
(c) SN; W_1, D_3



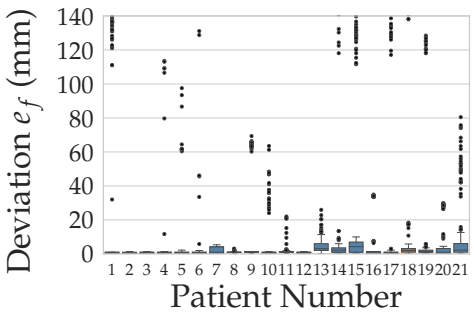
(d) DRN; W_1, D_3



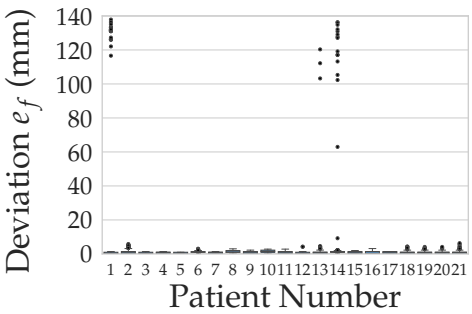
(e) SN; W_1, D_4



(f) DRN; W_1, D_4

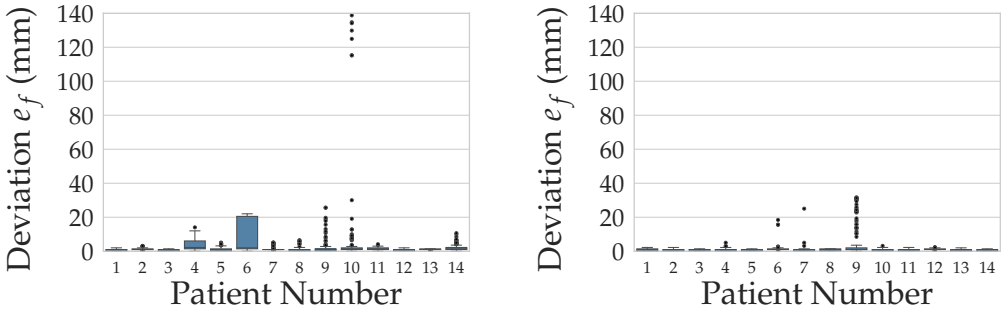


(g) SN; W_1, D_5



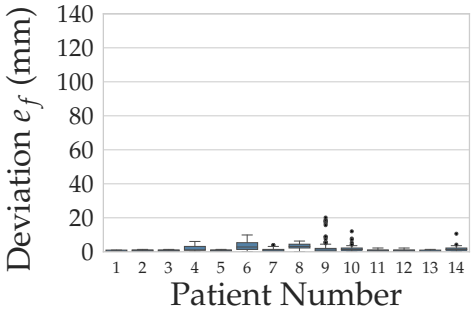
(h) DRN; W_1, D_5

Figure A.8: Robustness results of the training data order seed variation D_i with a fixed weight initialization seed W_1 of the standard (SNs) and the domain randomized networks (DRNs) on the AP angulation range.

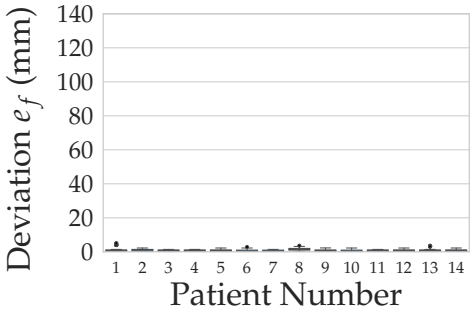


(a) SN; W_2, D_1

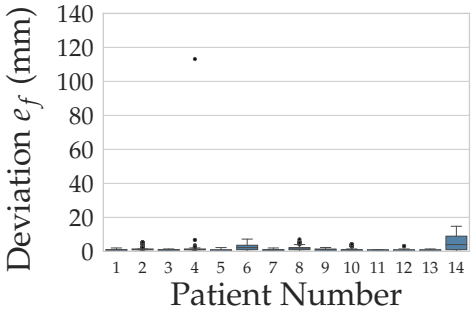
(b) DRN; W_2, D_1



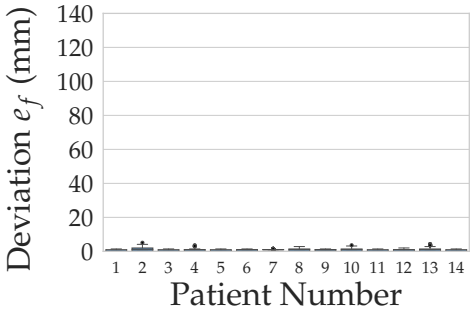
(c) SN; W_3, D_1



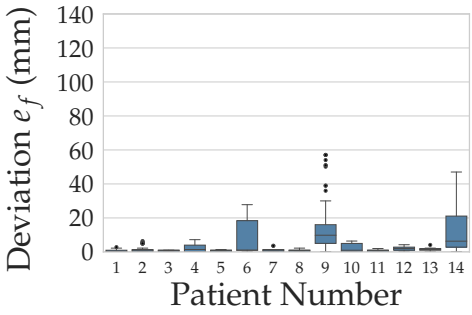
(d) DRN; W_3, D_1



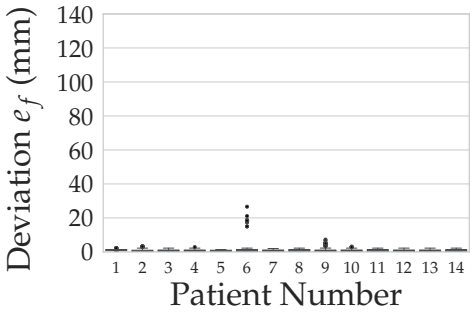
(e) SN; W_4, D_1



(f) DRN; W_4, D_1

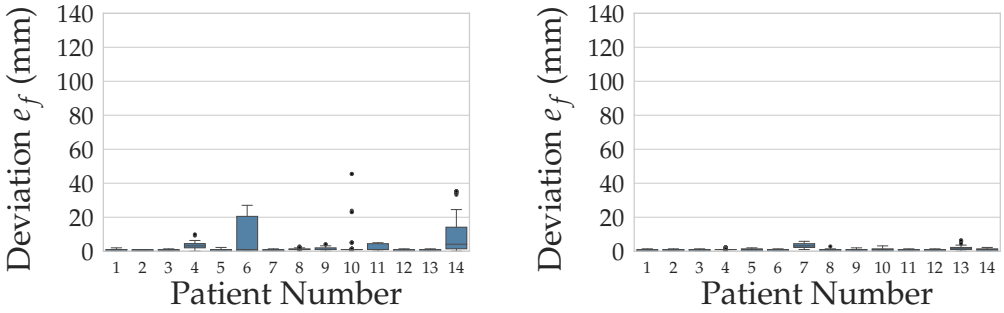


(g) SN; W_5, D_1



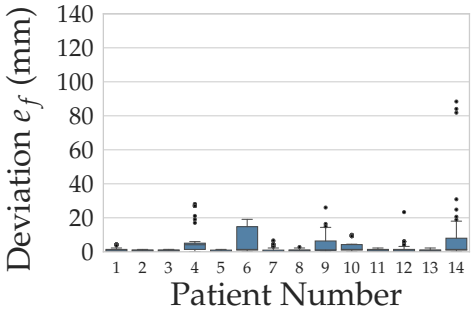
(h) DRN; W_5, D_1

Figure A.9: Robustness results of the weight initialization seed variation W_i with a fixed training data order seed D_1 of the standard (SNs) and the domain randomized networks (DRNs) on the RAO angulation range.

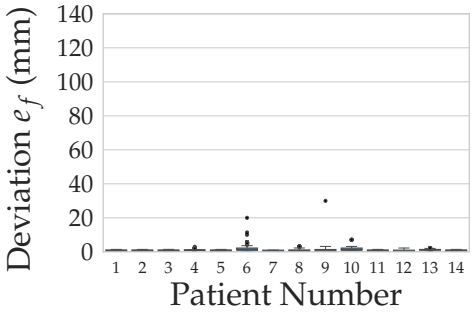


(a) SN; W_1, D_2

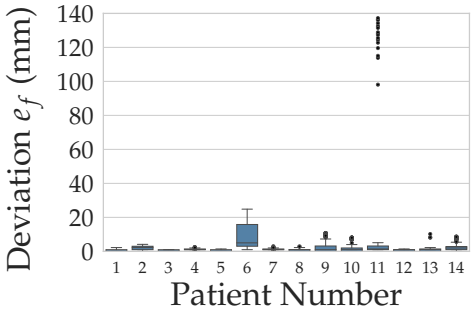
(b) DRN; W_1, D_2



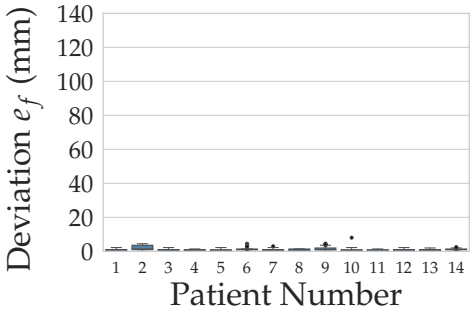
(c) SN; W_1, D_3



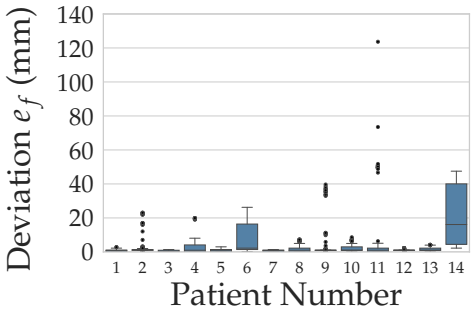
(d) DRN; W_1, D_3



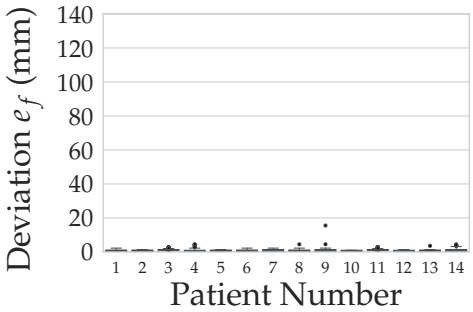
(e) SN; W_1, D_4



(f) DRN; W_1, D_4

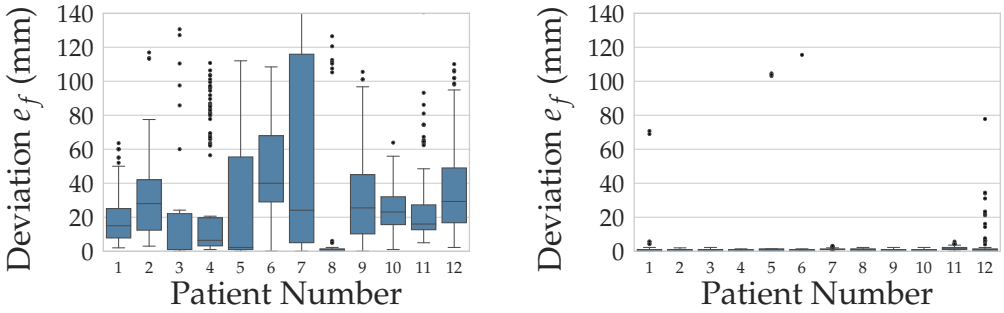


(g) SN; W_1, D_5



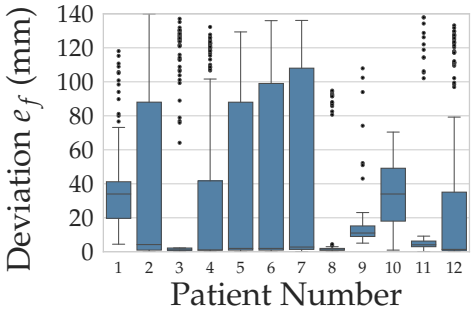
(h) DRN; W_1, D_5

Figure A.10: Robustness results of the training data order seed variation D_i with a fixed weight initialization seed W_1 of the standard (SNs) and the domain randomized networks (DRNs) on the RAO angulation range.

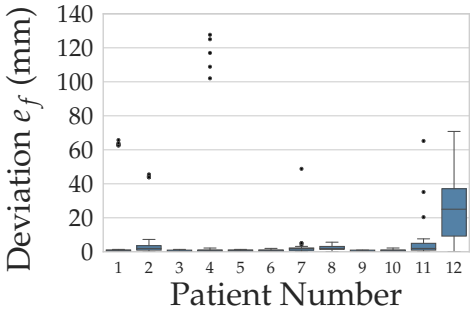


(a) SN; W_2, D_1

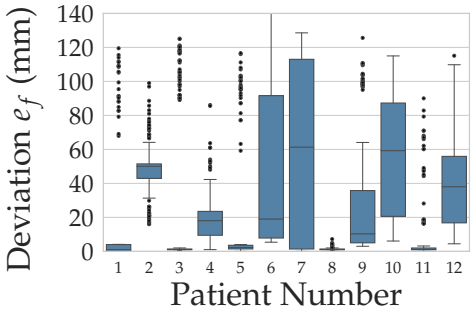
(b) DRN; W_2, D_1



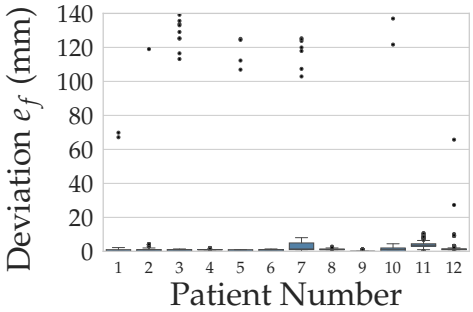
(c) SN; W_3, D_1



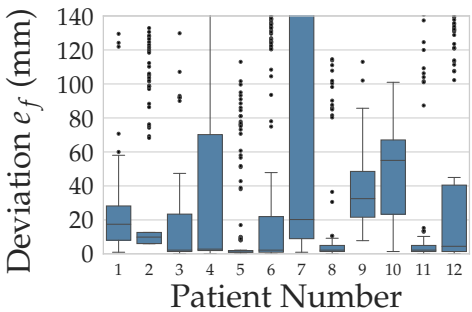
(d) DRN; W_3, D_1



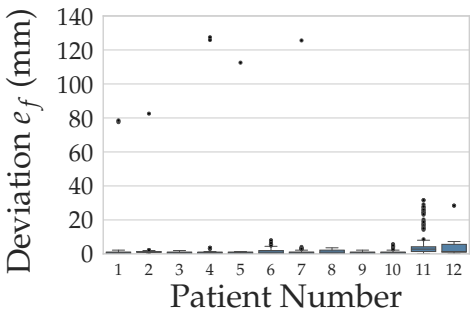
(e) SN; W_4, D_1



(f) DRN; W_4, D_1

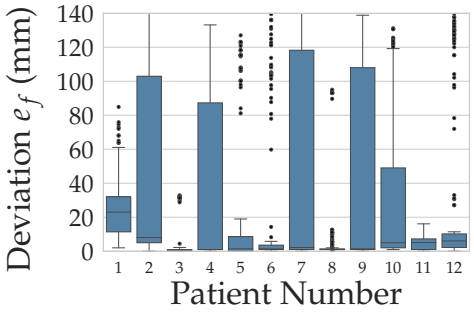


(g) SN; W_5, D_1

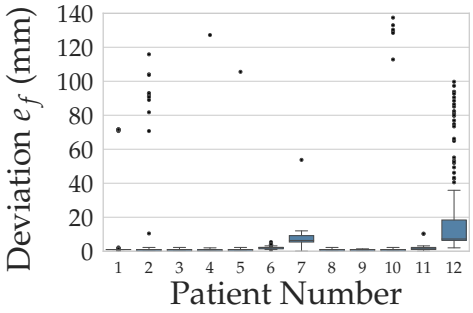


(h) DRN; W_5, D_1

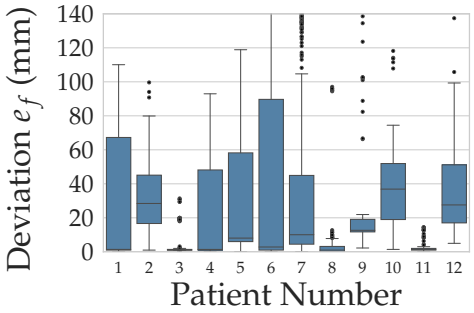
Figure A.11: Robustness results of the weight initialization seed variation W_i with a fixed training data order seed D_1 of the standard (SNs) and the domain randomized networks (DRNs) on the LAO angulation range.



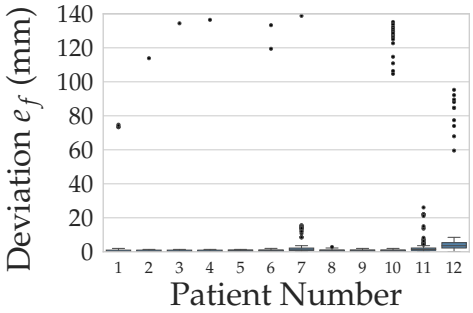
(a) SN; W_1, D_2



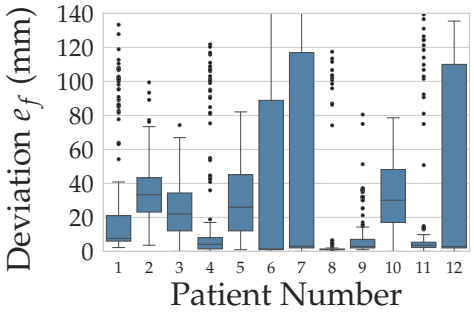
(b) DRN; W_1, D_2



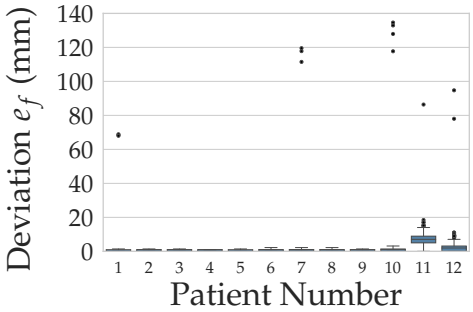
(c) SN; W_1, D_3



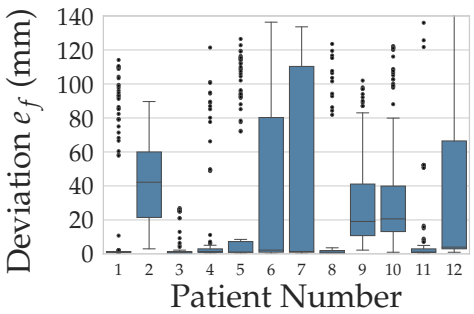
(d) DRN; W_1, D_3



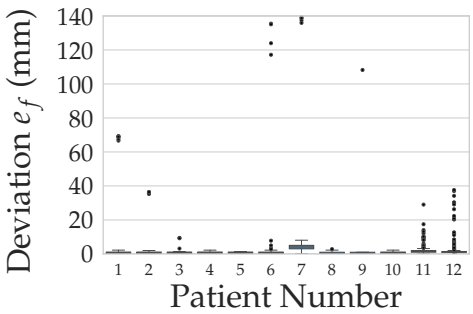
(e) SN; W_1, D_4



(f) DRN; W_1, D_4



(g) SN; W_1, D_5



(h) DRN; W_1, D_5

Figure A.12: Robustness results of the training data order seed variation D_i with a fixed weight initialization seed W_1 of the standard (SNs) and the domain randomized networks (DRNs) on the LAO angulation range.

Bibliography

[1] S. Claridge, Z. Chen, T. Jackson, K. De Silva, J. Behar, M. Sohal, J. Webb, E. Hyde, M. Lumley, K. Asrress, R. Williams, J. Bostock, M. Ali, J. Gill, M. O'Neill, R. Razavi, S. Niederer, D. Perera, and C. A. Rinaldi. "Effects of Epicardial and Endocardial Cardiac Resynchronization Therapy on Coronary Flow: Insights From Wave Intensity Analysis." In: *Journal of the American Heart Association* 4.12 (2015), pp. 1–12. DOI: 10.1161/JAHA.115.002626.

[2] J.-C. Daubert et al. "2012 EHRA / HRS Expert Consensus Statement on Cardiac Resynchronization Therapy in Heart Failure: Implant and Follow-up Recommendations and Management". In: *Heart Rhythm* 9.9 (2012), pp. 1524–1576. DOI: 10.1016/j.hrthm.2012.07.025.

[3] K. Mischke, C. Knackstedt, G. Mühlenbruch, T. Schimpf, P. Neef, M. Zarse, J. Plisiene, S. Stanzel, C. Eickholt, K. Fache, D. Frechen, E. Spüntrup, P. Hanrath, M. Kelm, and P. Schauerte. "Imaging of the coronary venous system: Retrograde coronary sinus angiography versus venous phase coronary angiograms". In: *International Journal of Cardiology* 119.3 (2007), pp. 339–343. DOI: 10.1016/j.ijcard.2006.07.148.

[4] P. Mountney, J. M. Behar, D. Toth, M. Panayiotou, S. Reiml, M. P. Jolly, R. Karim, L. Zhang, A. Brost, C. A. Rinaldi, and K. Rhode. "A Planning and Guidance Platform for Cardiac Resynchronization Therapy". In: *IEEE Transactions on Medical Imaging* 36.11 (2017), pp. 2366–2375. DOI: 10.1109/TMI.2017.2720158.

[5] J. M. Fitzpatrick. "Fiducial registration error and target registration error are uncorrelated". In: *Proceedings of SPIE* 7261 (2009), pp. 726102–726102–12. DOI: 10.1117/12.813601.

[6] M. V. N. Truong, A. Aslam, C. A. Rinaldi, R. Razavi, G. P. Penney, and K. S. Rhode. "Preliminary Investigation : 2D-3D Registration of MR and X-ray Cardiac Images Using Catheter Constraints". In: *MICCAI Workshop on Cardiovascular Interventional Imaging and Biophysical Modelling*. London, 2009, pp. 1–9.

[7] A. Brost, F. Bourier, L. Yatziv, M. Koch, J. Hornegger, N. Strobel, and K. Kurzidim. "First Steps towards Initial Registration for Electrophysiology Procedures". In: *SPIE Medical Imaging* 7964 (2011), 79641P–79641P–8. DOI: 10.1117/12.877654.

[8] J. Choi, P. Radau, R. Xu, and G. A. Wright. "X-Ray and Magnetic Resonance Imaging Fusion for Cardiac Resynchronization Therapy". In: *Medical Image Analysis* 31 (2016), pp. 98–107. DOI: 10.1016/j.media.2016.03.004.

[9] Y. Ma, S. Duckett, P. Chinchapatnam, G. Gao, A. Shetty, C. A. Rinaldi, T. Schaeffter, and K. S. Rhode. "MRI to X-ray Fluoroscopy Overlay for Guidance of Cardiac Resynchronization Therapy Procedures". In: *2010 Computers in Cardiology* (2010), pp. 229–232.

[10] T. L. Faber, C. A. Santana, E. V. Garcia, J. Candell-Riera, R. D. Folks, J. W. Peifer, A. Hopper, S. Aguade, J. Angel, and J. L. Klein. "Three-Dimensional Fusion of Coronary Arteries with Myocardial Perfusion Distributions: Clinical Validation". In: *Journal of Nuclear Medicine* 45.5 (2004), pp. 745–753.

[11] W. Zhou, X. Hou, M. Piccinelli, X. Tang, L. Tang, K. Cao, E. V. Garcia, J. Zou, and J. Chen. "3D Fusion of LV Venous Anatomy on Fluoroscopy Venograms With Epicardial Surface on SPECT Myocardial Perfusion Images for Guiding CRT LV Lead Placement". In: *JACC: Cardiovascular Imaging* 7.12 (2014), pp. 1239–1248. DOI: 10.1016/j.jcmg.2014.09.002.

[12] R. H. Anderson, R. Razavi, and A. M. Taylor. "Cardiac anatomy revisited". In: *Journal of Anatomy* 205.3 (2004), pp. 159–177. DOI: 10.1111/j.0021-8782.2004.00330.x.

[13] J. L. Titus. "Normal Anatomy of the Human Cardiac Conduction System". In: *Anesthesia & Analgesia* 52.4 (1973), pp. 508–517.

[14] J. P. Singh, S. Houser, E. K. Heist, and J. N. Ruskin. "The Coronary Venous Anatomy: A Segmental Approach to Aid Cardiac Resynchronization Therapy". In: *Journal of the American College of Cardiology* 46.1 (2005), pp. 68–74. DOI: 10.1016/j.jacc.2005.04.017.

[15] M. R. Bristow, L. A. Saxon, J. Boehmer, S. Krueger, D. A. Kass, T. De Marco, P. Carson, L. DiCarlo, D. DeMets, B. G. White, D. W. DeVries, and A. M. Feldman. "Cardiac-Resynchronization Therapy with or without an Implantable Defibrillator in Advanced Chronic Heart Failure". In: *New England Journal of Medicine* 350.21 (2004), pp. 2140–2150. DOI: 10.1056/NEJMoa032423.

[16] J. G. Cleland, J. C. Daubert, E. Erdmann, N. Freemantle, D. Gras, L. Kappenberger, and L. Tavazzi. "Longer-term effects of cardiac resynchronization therapy on mortality in heart failure [the CArdiac REsynchronization-Heart Failure (CARE-HF) trial extension phase]". In: *European Heart Journal* 27.16 (2006), pp. 1928–1932. DOI: 10.1093/eurheartj/ehl099.

[17] C. Daubert, M. R. Gold, W. T. Abraham, S. Ghio, C. Hassager, G. Goode, T. Szili-Török, and C. Linde. "Prevention of Disease Progression by Cardiac Resynchronization Therapy in Patients With Asymptomatic or Mildly Symptomatic Left Ventricular Dysfunction. Insights From the European Cohort of the REVERSE (Resynchronization Reverses Remodeling in Systolic Le". In: *Journal of the American College of Cardiology* 54.20 (2009), pp. 1837–1846. DOI: 10.1016/j.jacc.2009.08.011.

[18] W. B. Kannel. "Incidence and epidemiology of heart failure". In: *Heart Failure Reviews* 5.2 (2000), pp. 167–173.

[19] S. Baldasseroni et al. "Left bundle-branch block is associated with increased 1-year sudden and total mortality rate in 5517 outpatients with congestive heart failure: A report from the Italian Network on Congestive Heart Failure". In: *American Heart Journal* 143.3 (2002), pp. 398–405. DOI: 10.1067/mhj.2002.121264.

[20] A. P. Davie, C. M. Francis, M. P. Love, L. Caruana, I. R. Starkey, T. R. D. Shaw, G. R. Sutherland, and J. J. V. McMurray. "Value of the electrocardiogram in identifying heart failure due to left ventricular systolic dysfunction". In: *British Medical Journal* 312.7025 (1996), pp. 222–222. DOI: 10.1136/bmj.312.7025.222.

[21] A. Kashani and S. S. Barold. "Significance of QRS complex duration in patients with heart failure". In: *Journal of the American College of Cardiology* 46.12 (2005), pp. 2183–2192. DOI: 10.1016/j.jacc.2005.01.071.

[22] G. B. Bleeker, M. J. Schalij, S. G. Molhoek, H. F. Verwey, E. R. Holman, E. Boersma, P. Steendijk, E. E. Van Der Wall, and J. J. Bax. "Relationship between QRS duration and left ventricular dyssynchrony in patients with end-stage heart failure". In: *Journal of Cardiovascular Electrophysiology* 15.5 (2004), pp. 544–549. DOI: 10.1046/j.1540-8167.2004.03604.x.

[23] R. A. Nishimura and A. J. Tajik. "Evaluation of diastolic filling of left ventricle in health and disease: Doppler echocardiography is the clinician's Rosetta Stone". In: *Journal of the American College of Cardiology* 30.1 (1997), pp. 8–18. DOI: 10.1016/S0735-1097(97)00144-7.

[24] M. Dolgin, ed. *Nomenclature and criteria for diagnosis of diseases of the heart and great vessels / the Criteria Committee of the New York Heart Association*. 9th. Boston, Mass: Little, Brown & co, 1994, pp. 253–256.

[25] S. A. Strickberger, J. Conti, E. G. Daoud, E. Havranek, M. R. Mehra, I. L. Piña, and J. Young. "Patient selection for cardiac resynchronization therapy: From the Council on Clinical Cardiology Subcommittee on Electrocardiography and Arrhythmias and the Quality of Care and Outcomes Research Interdisciplinary Working Group, in collaboration with the H". In: *Circulation* 111.16 (2005), pp. 2146–2150. DOI: 10.1161/01.CIR.0000161276.09685.4A.

[26] G. B. Bleeker, M. J. Schalij, E. R. Holman, P. Steendijk, E. E. van der Wall, and J. J. Bax. "Cardiac Resynchronization Therapy in Patients With Systolic Left Ventricular Dysfunction and Symptoms of Mild Heart Failure Secondary to Ischemic or Nonischemic Cardiomyopathy". In: *American Journal of Cardiology* 98.2 (2006), pp. 230–235. DOI: 10.1016/j.amjcard.2006.01.080.

[27] A. Auricchio, H. Klein, B. Tockman, S. Sack, C. Stellbrink, J. Neuzner, A. Kramer, J. Ding, T. Pochet, A. Maarse, and J. Spinelli. "Transvenous biventricular pacing for heart failure: Can the obstacles be overcome". In: *American Journal of Cardiology* 83.5 B (1999). DOI: 10.1016/S0002-9149(98)01015-7.

[28] MedtronicCardiac. *What is Cardiac Resynchronization Therapy CRT, and how does it work?* https://www.youtube.com/watch?v=DdDPVTXX_iQ. 2014.

[29] C. Ypenburg, R. J. van Bommel, C. J. W. Borleffs, G. B. Bleeker, E. Boersma, M. J. Schalij, and J. J. Bax. "Long-Term Prognosis After Cardiac Resynchronization Therapy Is Related to the Extent of Left Ventricular Reverse Remodeling at Midterm Follow-Up". In: *Journal of the American College of Cardiology* 53.6 (2009), pp. 483–490. DOI: 10.1016/j.jacc.2008.10.032.

[30] N. Derval et al. "Optimizing Hemodynamics in Heart Failure Patients by Systematic Screening of Left Ventricular Pacing Sites. The Lateral Left Ventricular Wall and the Coronary Sinus Are Rarely the Best Sites". In: *Journal of the American College of Cardiology* 55.6 (2010), pp. 566–575. DOI: 10.1016/j.jacc.2009.08.045.

[31] C. Ypenburg, M. J. Schalij, G. B. Bleeker, P. Steendijk, E. Boersma, P. Dibbets-Schneider, M. P. Stokkel, E. E. Van Der Wall, and J. J. Bax. "Impact of viability and scar tissue on response to cardiac resynchronization therapy in ischaemic heart failure patients". In: *European Heart Journal* 28.1 (2007), pp. 33–41. DOI: 10.1093/eurheartj/ehl379.

[32] F. Leyva. "Cardiac resynchronization therapy guided by late gadolinium-enhancement cardiovascular magnetic resonance". In: *Journal of Cardiovascular Magnetic Resonance* 13.1 (2011), pp. 1–22.

[33] C. Ypenburg, R. J. van Bommel, V. Delgado, S. A. Mollema, G. B. Bleeker, E. Boersma, M. J. Schalij, and J. J. Bax. "Optimal Left Ventricular Lead Position Predicts Reverse Remodeling and Survival After Cardiac Resynchronization Therapy". In: *Journal of the American College of Cardiology* 52.17 (2008), pp. 1402–1409. DOI: 10.1016/j.jacc.2008.06.046.

[34] R. T. Murphy, G. Sigurdsson, S. Mulamalla, D. Agler, Z. B. Popovic, R. C. Starling, B. L. Wilkoff, J. D. Thomas, and R. A. Grimm. "Tissue Synchronization Imaging and Optimal Left Ventricular Pacing Site in Cardiac Resynchronization Therapy". In: *American Journal of Cardiology* 97.11 (2006), pp. 1615–1621. DOI: 10.1016/j.amjcard.2005.12.054.

[35] J. P. Singh, H. U. Klein, D. T. Huang, S. Reek, M. Kuniss, A. Quesada, A. Barsheshet, D. Cannom, I. Goldenberg, S. McNitt, J. P. Daubert, W. Zareba, and A. J. Moss. "Left ventricular lead position and clinical outcome in the multicenter automatic defibrillator implantation trial-cardiac resynchronization therapy (MADIT-CRT) trial". In: *Circulation* 123.11 (2011), pp. 1159–1166. DOI: 10.1161/CIRCULATIONAHA.110.000646.

[36] C. Thébault, E. Donal, C. Meunier, R. Gervais, B. Gerritse, M. R. Gold, W. T. Abraham, C. Linde, and J. C. Daubert. "Sites of left and right ventricular lead implantation and response to cardiac resynchronization therapy observations from the REVERSE trial". In: *European Heart Journal* 33.21 (2012), pp. 2662–2671. DOI: 10.1093/eurheartj/ehr505.

[37] E. K. Heist, D. Fan, T. Mela, D. Arzola-Castaner, V. Y. Reddy, M. Mansour, M. H. Picard, J. N. Ruskin, and J. P. Singh. "Radiographic left ventricular-right ventricular interlead distance predicts the acute hemodynamic response to cardiac resynchronization therapy". In: *American Journal of Cardiology* 96.5 (2005), pp. 685–690. DOI: 10.1016/j.amjcard.2005.04.045.

[38] M. Biffi, D. V. Exner, G. H. Crossley, B. Ramza, B. Coutu, G. Tomassoni, W. Kranig, S. Li, N. Kristiansen, and F. Voss. "Occurrence of phrenic nerve stimulation in cardiac resynchronization therapy patients: The role of left ventricular lead type and placement site". In: *Europace* 15.1 (2013), pp. 77–82. DOI: 10.1093/europace/eus237.

[39] S. Goetze, P. Defaye, A. Bauer, M. Merkel, O. Bizeau, S. Treusch, K. Contzen, C. Juenger, and J. Winter. "Phrenic nerve stimulation in CRT patients and benefits of electronic lead repositioning: The ERACE trial". In: *Journal of Interventional Cardiac Electrophysiology* 38.1 (2013), pp. 1–9. DOI: 10.1007/s10840-013-9811-9.

[40] G. Moubarak, A. Bouzeman, J. Ollitrault, F. Anselme, and S. Cazeau. "Phrenic nerve stimulation in cardiac resynchronization therapy". In: *Journal of Interventional Cardiac Electrophysiology* 41.1 (2014), pp. 15–21. DOI: 10.1007/s10840-014-9917-8.

[41] E. Thomas, D. Toth, T. Kurzendorfer, K. Rhode, and P. Mountney. "Mechanical Activation Computation from Fluoroscopy for Guided Cardiac Resynchronization Therapy". In: *40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. Honolulu, HI, USA, 2018, pp. 592–595. DOI: 10.1109/EMBC.2018.8512434.

[42] F. Z. Khan, M. S. Virdee, C. R. Palmer, P. J. Pugh, D. O'Halloran, M. Elsik, P. A. Read, D. Begley, S. P. Fynn, and D. P. Dutka. "Targeted left ventricular lead placement to guide cardiac resynchronization therapy: The TARGET study: A randomized, controlled trial". In: *Journal of the American College of Cardiology* 59.17 (2012), pp. 1509–1518. DOI: 10.1016/j.jacc.2011.12.030. arXiv: 0804.4698.

[43] A. Sommer, M. B. Kronborg, B. L. Norgaard, S. H. Poulsen, K. Bouchelouche, M. Böttcher, H. K. Jensen, J. M. Jensen, J. Kristensen, C. Gerdes, P. T. Mortensen, and J. C. Nielsen. "Multimodality imaging-guided left ventricular lead placement improves clinical outcome in cardiac resynchronization therapy: A randomized controlled trial". In: *European Heart Journal* 36.Lv (2015), p. 372. DOI: 10.1093/eurheartj/ehv399.

[44] B. J. Sieniewicz, J. Gould, B. Porter, B. S. Sidhu, J. M. Behar, S. Claridge, S. Niederer, and C. A. Rinaldi. "Optimal site selection and image fusion guidance technology to facilitate cardiac resynchronization therapy". In: *Expert Review of Medical Devices* 15.8 (2018), pp. 555–570. DOI: 10.1080/17434440.2018.1502084.

[45] E. Adelstein, M. B. Alam, D. Schwartzman, S. Jain, J. Marek, J. Gorcsan, and S. Saba. "Effect of echocardiography-guided left ventricular lead placement for cardiac resynchronization therapy on mortality and risk of defibrillator therapy for ventricular arrhythmias in heart failure patients (from the Speckle Tracking Assisted Resynchronizat". In: *American Journal of Cardiology* 113.9 (2014), pp. 1518–1522. DOI: 10.1016/j.amjcard.2014.01.431.

[46] S. Saba, J. Marek, D. Schwartzman, S. Jain, E. Adelstein, P. White, O. A. Oyenuga, T. Onishi, P. Soman, and J. Gorcsan. "Echocardiography-guided left ventricular lead placement for cardiac resynchronization therapy results of the speckle tracking assisted resynchronization therapy for electrode region trial". In: *Circulation: Heart Failure* 6.3 (2013), pp. 427–434. DOI: 10.1161/CIRCHEARTFAILURE.112.000078.

[47] S. Saksena, A. M. Simon, P. Mathew, and R. Nagarakanti. "Intracardiac echocardiography-guided cardiac resynchronization therapy: Technique and clinical application". In: *PACE - Pacing and Clinical Electrophysiology* 32.8 (2009), pp. 1030–1039. DOI: 10.1111/j.1540-8159.2009.02435.x.

[48] R. J. Kim, E. Wu, A. Rafael, E.-L. Chen, M. A. Parker, O. Simonetti, F. J. Klocke, R. O. Bonow, and R. M. Judd. "The use of contrast-enhanced magnetic resonance imaging to identify reversible myocardial dysfunction".

In: *The New England Journal of Medicine* 343.20 (2000), pp. 1445–1453. DOI: 10.1056/NEJM200011163432003.

[49] J. A. White, R. Yee, X. Yuan, A. Krahn, A. Skanes, M. Parker, G. Klein, and M. Drangova. "Delayed Enhancement Magnetic Resonance Imaging Predicts Response to Cardiac Resynchronization Therapy in Patients With Intraventricular Dyssynchrony". In: *Journal of the American College of Cardiology* 48.10 (2006), pp. 1953–1960. DOI: 10.1016/j.jacc.2006.07.046.

[50] A. K. Shetty, S. G. Duckett, M. R. Ginks, Y. Ma, M. Sohal, J. Bostock, S. Kapetanakis, J. P. Singh, K. Rhode, M. Wright, M. D. O'Neill, J. S. Gill, G. Carr-White, R. Razavi, and C. A. Rinaldi. "Cardiac magnetic resonance-derived anatomy, scar, and dyssynchrony fused with fluoroscopy to guide LV lead placement in cardiac resynchronization therapy: A comparison with acute haemodynamic measures and echocardiographic reverse remodelling". In: *European Heart Journal Cardiovascular Imaging* 14.7 (2013), pp. 692–699. DOI: 10.1093/ehjci/jes270.

[51] J. M. Behar, R. Rajani, A. Pourmorteza, R. Preston, O. Razeghi, S. Niederer, S. Adhya, S. Claridge, T. Jackson, B. Sieniewicz, J. Gould, G. Carr-White, R. Razavi, E. McVeigh, and C. A. Rinaldi. "Comprehensive use of cardiac computed tomography to guide left ventricular lead placement in cardiac resynchronization therapy". In: *Heart Rhythm* 14.9 (2017), pp. 1364–1372. DOI: 10.1016/j.hrthm.2017.04.041.

[52] S. Lehner, C. Uebleis, F. Schüßler, A. Haug, S. Kääb, P. Bartenstein, S. D. Van Kriekinge, G. Germano, H. Estner, and M. Hacker. "The amount of viable and dyssynchronous myocardium is associated with response to cardiac resynchronization therapy: Initial clinical results using multiparametric ECG-gated [18F]FDG PET". In: *European Journal of Nuclear Medicine and*

Molecular Imaging 40.12 (2013), pp. 1876–1883. DOI: 10.1007/s00259-013-2516-6.

[53] R. Sciagra, M. Giaccardi, M. C. Porciani, A. Colella, A. Michelucci, P. Pieragnoli, G. Gensini, A. Pupi, and L. Padeletti. "Myocardial perfusion imaging using gated SPECT in heart failure patients undergoing cardiac resynchronization therapy". In: *Journal of Nuclear Medicine* 45.2 (2004), pp. 164–168.

[54] D. Tomaževič, B. Likar, T. Slivnik, and F. Pernuš. "3-D / 2-D Registration of CT and MR to X-Ray Images". In: *IEEE Transactions on Medical Imaging* 22.11 (2003), pp. 1407–1416. DOI: 10.1109/TMI.2003.819277.

[55] S. Miao, S. Piat, P. Fischer, A. Tuysuzoglu, P. Mewes, T. Mansi, and R. Liao. "Dilated FCN for Multi-Agent 2D / 3D Medical Image Registration". In: *Association for the Advancement of Artificial Intelligence* (2018).

[56] M. D. Cerqueira, N. J. Weissman, V. Dilsizian, A. K. Jacobs, S. Kaul, W. K. Laskey, D. J. Pennell, J. A. Rumberger, T. Ryan, and M. S. Verani. "Standardized Myocardial Segmentation and Nomenclature for Tomographic Imaging of the Heart: A Statement for Healthcare Professionals From the Cardiac Imaging". In: *Circulation* 105.4 (2002), pp. 539–542. DOI: 10.1161/hc0402.102975.

[57] J. M. Behar, H. M. S. Chin, S. Fearn, J. O. Ormerod, J. Gamble, P. W. Foley, J. Bostock, S. Claridge, T. Jackson, M. Sohal, A. P. Antoniadis, R. Razavi, T. R. Betts, N. Herring, and C. A. Rinaldi. "Cost-Effectiveness Analysis of Quadripolar Versus Bipolar Left Ventricular Leads for Cardiac Resynchronization Defibrillator Therapy in a Large, Multicenter UK Registry". In: *JACC: Clinical Electrophysiology* 3.2 (2017), pp. 107–116. DOI: 10.1016/j.jacep.2016.04.009.

[58] A. P. Antoniadis, J. M. Behar, S. Claridge, T. Jackson, M. Sohal, and C. A. Rinaldi. "Multisite Pacing for Cardiac Resynchronization Therapy: Promise and Pitfalls". In: *Current Cardiology Reports* 18.7 (2016), pp. 1–6. DOI: 10.1007/s11886-016-0741-x.

[59] J. M. Behar, J. Bostock, A. P. Zhu Li, H. M. S. Chin, S. Jubb, E. Lent, J. Gamble, P. W. Foley, T. R. Betts, C. A. Rinaldi, and N. Herring. "Cardiac resynchronization therapy delivered via a multipolar left ventricular lead is associated with reduced mortality and elimination of phrenic nerve stimulation: Long-term follow-up from a multicenter registry". In: *Journal of Cardiovascular Electrophysiology* 26.5 (2015), pp. 540–546. DOI: 10.1111/jce.12625.

[60] T. Jackson, R. Lenarczyk, M. Sterlinski, A. Sokal, D. Francis, Z. Whinnett, F. Van Heuverswyn, M. Vanderheyden, J. Heynens, B. Stegemann, R. Cornelussen, and C. A. Rinaldi. "Left ventricular scar and the acute hemodynamic effects of multivein and multipolar pacing in cardiac resynchronization". In: *IJC Heart and Vasculature* 19 (2018), pp. 14–19. DOI: 10.1016/j.ijcha.2018.03.006.

[61] V. Kutyifa, B. Merkely, S. Szilágyi, E. Zima, A. Róka, Á. Király, I. Osztheimer, L. Molnár, G. Széplaki, and L. Gellér. "Usefulness of electroanatomical mapping during transseptal endocardial left ventricular lead implantation". In: *Europace* 14.4 (2012), pp. 599–604. DOI: 10.1093/europace/eur353.

[62] B. J. Sieniewicz, T. Jackson, S. Claridge, H. Pereira, J. Gould, B. Sidhu, B. Porter, S. Niederer, C. Yao, and C. A. Rinaldi. "Optimization of CRT programming using non-invasive electrocardiographic imaging to assess the acute electrical effects of multipoint pacing". In: *Journal of Arrhythmia* July 2018 (2019), pp. 1–9. DOI: 10.1002/joa3.12153.

[63] J. M. Behar, T. Jackson, E. Hyde, S. Claridge, J. Gill, J. Bostock, M. Sohal, B. Porter, M. O'Neill, R. Razavi, S. Niederer, and C. A. Rinaldi. "Optimized Left Ventricular Endocardial Stimulation Is Superior to Optimized Epicardial Stimulation in Ischemic Patients With Poor Response to Cardiac Resynchronization Therapy: A Combined Magnetic Resonance Imaging, Electroanatomic Contact Mapping, and He". In: *JACC: Clinical Electrophysiology* 2.7 (2016), pp. 799–809. DOI: 10.1016/j.jacep.2016.04.006.

[64] A. Auricchio, P. P. Delnoy, C. Butter, J. Brachmann, L. Van Erven, S. Spitzer, T. Moccetti, M. Seifert, T. Markou, K. Laszo, and F. Regoli. "Feasibility, safety, and short-term outcome of leadless ultrasound-based endocardial left ventricular resynchronization in heart failure patients: Results of the Wireless Stimulation Endocardially for CRT (WiSE-CRT) study". In: *Europace* 16.5 (2014), pp. 681–688. DOI: 10.1093/europace/eut435.

[65] V. Gorbunova, P. Lo, H. Ashraf, A. Dirksen, M. Nielsen, and M. De Bruijne. "Weight preserving image registration for monitoring disease progression in lung CT". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 5242 LNCS. PART 2. 2008, pp. 863–870. DOI: 10.1007/978-3-540-85990-1-104.

[66] D. Toth, M. Pfister, A. Maier, M. Kowarschik, and J. Hornegger. "Adaption of 3D Models to 2D X-Ray Images during Endovascular Abdominal Aneurysm Repair". In: *Medical Image Computing and Computer-Assisted Intervention — MICCAI 2015*. Vol. 9349. Munich: Springer, 2015, pp. 339–346. DOI: 10.1007/978-3-319-24574-4.

[67] C. R. Maurer and J. M. Fitzpatrick. "A Review of Medical Image Registration". In: *Interactive Image-Guided Neurosurgery*. Park Ridge, IL: American Association of Neurological Surgeons, 1993, pp. 17–44.

[68] J. B. A. Maintz and M. A. Viergever. "A Survey of Medical Image Registration". In: *Medical Image Analysis* 2.1 (1998), pp. 1–36. DOI: 10.1.1.39.4417.

[69] D. L. G. Hill, P. G. Batchelor, M. Holden, and D. J. Hawkes. "Medical Image Registration". In: *Physics in Medicine and Biology* 46.3 (2001), R1–R45. DOI: 10.1088/0031-9155/46/3/201.

[70] J. P. Pluim and J. M. Fitzpatrick. "Image Registration". In: *IEEE Transactions on Medical Imaging* 22.11 (2012), pp. 2002–2004. DOI: 10.1007/978-1-4471-2458-0.

[71] R. Liao, L. Zhang, Y. Sun, S. Miao, and C. Chefd'Hotel. "A Review of recent advances in registration techniques applied to minimally invasive therapy". In: *IEEE Transactions on Multimedia* 15.5 (2013), pp. 983–1000. DOI: 10.1109/TMM.2013.2244869.

[72] L. Tang and G. Hamarneh. "Medical Image Registration: A Review". In: *Medical Imaging: Technology and Applications*. 2013, pp. 619–660. DOI: 10.1109/MEMB.2002.1175169.

[73] F. P. M. Oliveira and J. M. R. S. Tavares. "Medical image registration: A review". In: *Computer Methods in Biomechanics and Biomedical Engineering* 17.2 (2014), pp. 73–93. DOI: 10.1080/10255842.2012.670855.

[74] M. A. Viergever, J. B. Maintz, S. Klein, K. Murphy, M. Staring, and J. P. Pluim. "A survey of medical image registration – under review". In: *Medical Image Analysis* 33 (2016), pp. 140–144. DOI: 10.1016/j.media.2016.06.030.

[75] B. B. Avants, C. L. Epstein, M. Grossman, and J. C. Gee. "Symmetric diffeomorphic image registration with cross-correlation: evaluating automated

labeling of elderly and neurodegenerative brain". In: *Medical Image Analysis* 12.1 (2008), pp. 26–41. DOI: 10.1016/j.media.2007.06.004. arXiv: NIHMS150003.

[76] W. M. Wells III., P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis. "Multi-Modal Volume Registration by Maximization of Mutual Information". In: *Medical Image Analysis* 1.1 (1996), pp. 35–51. DOI: 10.1016/S1361-8415(96)80004-1.

[77] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. "Multimodality Image Registration by Maximization of Mutual Information". In: *IEEE Transactions on Medical Imaging* 16.2 (1997), pp. 187–198. DOI: 10.1109/42.563664.

[78] J. P. Pluim, J. B. A. Maintz, and M. A. Viergever. "Mutual-Information-Based Registration of Medical Images: A Survey". In: *IEEE Transactions on Medical Imaging* 22.8 (2003), pp. 986–1004. DOI: 10.1063/1.468183.

[79] D. Lowe. "Object Recognition fromLocal Scale-Invariant Features". In: *IEEE International Conference on Computer Vision*. Corfu, Greece, 1999, pp. 1150–1157. DOI: 10.1109/ICCV.1999.790410. arXiv: 0112017 [cs].

[80] C. M. Kramer, J. Barkhausen, S. D. Flamm, R. J. Kim, and E. Nagel. "Standardized cardiovascular magnetic resonance imaging (CMR) protocols, society for cardiovascular magnetic resonance: board of trustees task force on standardized protocols". In: *Journal of Cardiovascular Magnetic Resonance* 10.35 (2008), pp. 1–10. DOI: 10.1186/1532-429X-10-Received.

[81] S. Nithiananthan, S. Schafer, A. Uneri, D. J. Mirota, J. W. Stayman, W. Zbijewski, K. K. Brock, M. J. Daly, H. Chan, J. C. Irish, and J. H. Siewerdsen. "Demons deformable registration of CT and cone-beam CT using an

iterative intensity matching approach". In: *Medical Physics* 38.4 (2011), pp. 1785–1798. DOI: 10.1118/1.3555037.

[82] F. L. Bookstein. "Principal warps: thin-plate splines and the decomposition of deformations". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11.6 (1989), pp. 567–585. DOI: 10.1109/34.24792.

[83] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes. "Nonrigid registration using free-form deformations: application to breast MR images." In: *IEEE Transactions on Medical Imaging* 18.8 (1999), 18:712–21.

[84] T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache. "Diffeomorphic demons: efficient non-parametric image registration." In: *NeuroImage* 45.1 Suppl (2009), S61–S72. DOI: 10.1016/j.neuroimage.2008.10.040.

[85] M. Holden. "A Review of Geometric Transformations for Nonrigid Body Registration". In: *IEEE Transactions on Medical Imaging* 27.1 (2008), pp. 2–18.

[86] O. M. Henriksen, M. N. Lonsdale, T. D. Jensen, K. L. Weikop, O. Holm, B. Duus, and L. Friberg. "Two-Dimensional Image Fusion of Planar Bone Scintigraphy and Radiographs in Patients with Clinical Scaphoid Fracture: An Imaging Study". In: *Acta Radiologica* 50.1 (2009), pp. 71–77. DOI: 10.1080/02841850802562089.

[87] J. Betancur, A. Simon, F. Tavard, B. Langella, C. Leclercq, and M. Garreau. "Segmentation-Free MRI to CT 3D Registration for Cardiac Resynchronization Therapy Optimization". In: *Computing in Cardiology* 39 (2012), pp. 701–704.

[88] P. Markelj, D. Tomaževič, B. Likar, and F. Pernuš. "A Review of 3D/2D Registration Methods for Image-Guided Interventions". In: *Medical Image Analysis* 16.3 (2012), pp. 642–661. DOI: 10.1016/j.media.2010.03.005.

[89] K. S. Rhode, D. L. G. Hill, P. J. Edwards, J. Hipwell, D. Rueckert, G. Sanchez-Ortiz, S. Hegde, V. Rahunathan, and R. Razavi. "Registration and Tracking to Integrate X-ray and MR Images in an XMR Facility". In: *IEEE Transactions on Medical Imaging* 22.11 (2003), pp. 1369–1378. DOI: 10.1109/TMI.2003.819275.

[90] G.-A. Turgeon, G. Lehmann, G. Guiraudon, M. Drangova, D. Holdsworth, and T. Peters. "2D-3D Registration of Coronary Angiograms for Cardiac Procedure Planning and Guidance". In: *Medical Physics* 32.12 (2005), pp. 3737–3749. DOI: 10.1118/1.2123350.

[91] A. Krizhevsky, I. Sutskever, and G. E. Hinton. "ImageNet Classification with Deep Convolutional Neural Networks". In: *Advances In Neural Information Processing Systems*. 2012, pp. 1–9. DOI: http://dx.doi.org/10.1016/j.protcy.2014.09.007. arXiv: 1102.0183.

[92] J. Long, E. Shelhamer, and T. Darrell. "Fully Convolutional Networks for Semantic Segmentation". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015, pp. 3431–3440. DOI: 10.1109/CVPR.2015.7298965. arXiv: 1503.06350.

[93] P. Wohlhart and V. Lepetit. "Learning Descriptors for Object Recognition and 3D Pose Estimation". In: 22.1 (2017), pp. 99–131. DOI: 10.1016/S0733-8619(03)00096-3. arXiv: 1707.04192.

[94] D. C. Ciresan, A. Giusti, L. M. Gambardella, and J. Schmidhuber. "Deep Neural Networks Segment Neuronal Membranes in Electron Microscopy

Images". In: *Advances in Neural Information Processing Systems*. 2012, 2843–2851.

[95] K. Fritscher, P. Raudaschl, P. Zaffino, M. F. Spadea, G. C. Sharp, and R. Schubert. "Deep Neural Networks for Fast Segmentation of 3D Medical Images". In: *Medical Image Computing and Computer-Assisted Intervention*. Athens, Greece: Springer, 2016, pp. 158–165. DOI: 10.1007/978-3-319-46720-7.

[96] J. Jiang, S. Zheng, A. W. Toga, and Z. Tu. "Learning Based Coarse-to-Finene Image Registration". In: *Computer Vision and Pattern Recognition*. 2008.

[97] G. Wu, M. Kim, Q. Wang, Y. Gao, S. Liao, and D. Shen. "Unsupervised Deep Feature Learning for Deformable Registration of MR Brain Images". In: *Medical Image Computing and Computer-Assisted Intervention*. Springer, 2013, pp. 649–656.

[98] X. Cao, Y. Gao, J. Yang, G. Wu, and D. Shen. "Learning-Based Multimodal Image Registration for Prostate Cancer Radiation Therapy". In: *Medical Image Computing and Computer-Assisted Intervention*. Athens, Greece: Springer, 2016. DOI: 10.1007/978-3-319-46720-7.

[99] C.-r. Chou, B. Frederick, G. Mageras, and S. Chang. "Robust and Fast 2D / 3D Image Registration using Regression Learning". In: *Computer Vision and Image Understanding* 117.9 (2013), pp. 1095–1106.

[100] B. Gutierrez-Becker, D. Mateus, L. Peter, and N. Navab. "Learning Optimization Updates for Multimodal Registration". In: *Medical Image Computing and Computer-Assisted Intervention*. Athens, Greece: Springer, 2016, pp. 19–27. DOI: 10.1007/978-3-319-46720-7.

[101] S. Miao, Z. J. Wang, Y. Zheng, and R. Liao. "Real-time 2D/3D registration via CNN regression". In: *Proceedings - International Symposium on Biomedical Imaging* 2016-June (2016), pp. 1430–1434. DOI: 10.1109/ISBI.2016.7493536. arXiv: 1507.07505.

[102] R. Liao, S. Miao, P. de Tournemire, S. Grbic, A. Kamen, T. Mansi, and D. Comaniciu. "An Artificial Agent for Robust Image Registration". In: (2016), pp. 4168–4175. DOI: 10.1103/PhysRevX.5.041024. arXiv: 1611.10336.

[103] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis. "Human-level control through deep reinforcement learning". In: *Nature* 518.7540 (2015), pp. 529–533. DOI: 10.1038/nature14236. arXiv: 1604.03986.

[104] J. Krebs, T. Mansi, H. Delingette, L. Zhang, F. C. Ghesu, S. Miao, A. K. Maier, N. Ayache, R. Liao, and A. Kamen. "Robust non-rigid registration through agent-based action learning". In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2017, pp. 344–352.

[105] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. "Dropout: A Simple Way to Prevent Neural Networks from Over tting". In: *Journal of Machine Learning Research* 15.1 (2014), pp. 1929–1958.

[106] T. Mäkelä, P. Clarysse, O. Sipilä, N. Pauna, Q. C. Pham, T. Katila, and I. E. Magnin. "A Review of Cardiac Image Registration Methods". In: *IEEE Transactions on Medical Imaging* 21.9 (2002), pp. 1011–1021.

[107] J. Sra. "Cardiac Image Registration". In: *Journal of Atrial Fibrillation* 1.3 (2008), pp. 145–160. DOI: 10.4022/jafib.v1i1.387.

[108] K. Rhode and M. Sermesant. "Modeling and Registration for Electrophysiology Procedures Based on Three-Dimensional Imaging". In: *Current Cardiovascular Imaging Reports* 4.2 (2011), pp. 116–126. DOI: 10.1007/s12410-011-9067-7.

[109] J. Wetzl, F. Lugauer, R. Kroeker, M. Schmidt, A. Maier, and C. Forman. "Free-Breathing Self-Navigated Isotropic 3-D CINE Imaging of the Whole Heart using Adaptive Triggering and Retrospective Gating". In: *Proceedings of the 25h Annual Meeting of ISMRM, Honolulu, HI, USA* (2017).

[110] K. S. Rhode, M. Sermesant, D. Brogan, S. Hegde, J. Hipwell, P. Lambiase, E. Rosenthal, C. Bucknall, S. A. Qureshi, J. S. Gill, R. Razavi, and D. L. G. Hill. "A System for Real-Time XMR Guided Cardiovascular Intervention". In: *IEEE Transactions on Medical Imaging* 24.11 (2005), pp. 1428–1440. DOI: 10.1109/tmi.2005.856731.

[111] K. S. Rhode, M. Sermesant, S. R. Hegde, G. I. Sánchez-Ortiz, D. Rueckert, R. Razavi, and D. L. G. Hill. "XMR guided cardiac electrophysiology study and radio frequency ablation". In: *SPIE Medical Imaging* 5369.0 (2004), pp. 10–21. DOI: 10.1117/12.535103.

[112] V. Rasche, M. Mansour, V. Reddy, J. P. Singh, A. Qureshi, R. Manzke, S. Sokka, and J. Ruskin. "Fusion of Three-Dimensional X-Ray Angiography and Three-Dimensional Echocardiography". In: *International Journal of Computer Assisted Radiology and Surgery* 2.5 (2008), pp. 293–303. DOI: 10.1007/s11548-007-0142-0.

[113] R. de Silva, L. F. Gutiérrez, A. N. Raval, E. R. McVeigh, C. Ozturk, and R. J. Lederman. "X-Ray Fused with Magnetic Resonance Imaging (XFM)

to Target Endomyocardial Injections: Validation in a Swine Model of Myocardial Infarction." In: *Circulation* 114.22 (2006), pp. 2342–2350. DOI: 10.1161/CIRCULATIONAHA.105.598524.

[114] M. V. N. Truong, A. Aslam, M. Ginks, C. A. Rinaldi, R. Rezavi, G. P. Penney, and K. S. Rhode. "2D-3D Registration of Cardiac Images using Catheter Constraints". In: *IEEE Computing in Cardiology* 36 (2009), pp. 605–608.

[115] M. Truong, T. Gordon, R. Razavi, G. Penney, and K. S. Rhode. "Analysis of Catheter-Based Registration with Vessel-Radius Weighting of 3D CT Data to 2D X-Ray for Cardiac Catheterisation Procedures in a Phantom Study". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 7085 LNCS (2012), pp. 139–148. DOI: 10.1007/978-3-642-28326-0_14.

[116] D. Rivest-Henault, H. Sundar, and M. Cheriet. "Nonrigid 2D/3D Registration of Coronary Artery Models with Live Fluoroscopy for Guidance of Cardiac Interventions". In: *IEEE Transactions on Medical Imaging* 31.8 (2012), pp. 1557–1572. DOI: 10.1109/TMI.2012.2195009.

[117] M. M. Khalil, J. L. Tremoleda, T. B. Bayomy, and W. Gsell. "Molecular SPECT Imaging: An Overview". In: *International Journal of Molecular Imaging* 2011 (2011), pp. 1–15. DOI: 10.1155/2011/796025.

[118] F. Tavard, A. Simon, C. Leclercq, P. Mabo, A. I. Hernandez, and M. Garreau. "Data Fusion of Left Ventricle Electro-Anatomic Mapping and Multislice Computerized Tomography for Cardiac Resynchronisation Therapy Optimization". In: *IEEE Computers in Cardiology*. 2009, pp. 1745–1748.

[119] F. Tavard, A. Simon, E. Donal, A. I. Hernandez, and M. Garreau. "Fusion of MSCT Imaging, Electro-Anatomical Mapping and Speckle Tracking Echocardiography for the Characterization of Local Electro-Mechanical

Delays in CRT Optimization". In: *IEEE Computing in Cardiology*. 2010, pp. 401–404.

[120] F. Tavard, A. Simon, C. Leclercq, E. Donal, A. I. Hernandez, and M. Garreau. "Multimodal registration and data fusion for cardiac resynchronization therapy optimization". In: *IEEE Transactions on Medical Imaging* 33.6 (2014), pp. 1363–1372. DOI: 10.1109/TMI.2014.2311694.

[121] Y. Zheng, B. Georgescu, A. Barbu, M. Scheuering, and D. Comaniciu. "Four-Chamber Heart Modeling and Automatic Segmentation for 3D Cardiac CT Volumes". In: *IEEE Transactions on Medical Imaging* 27.11 (2008), pp. 691416–691416–12. DOI: 10.1117/12.770710.

[122] D. L. G. Hill and D. J. Hawkes. "Medical Image Registration Using Knowledge of Adjacency of Anatomical Structures". In: *Image and Vision Computing* 12.3 (1994), pp. 173–178. DOI: 10.5244/C.7.44.

[123] P. Besl and N. McKay. *A Method for Registration of 3-D Shapes*. 1992. DOI: 10.1109/34.121791.

[124] A. Myronenko and X. Song. "Point-Set Registration: Coherent Point Drift". In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 32.12 (2010), pp. 2262–2275. DOI: 10.1109/TPAMI.2010.46. arXiv: 0905.2635.

[125] J. Yang, H. Li, and Y. Jia. "Go-ICP: Solving 3D Registration Efficiently and Globally Optimally". In: *2013 IEEE International Conference on Computer Vision*. 2013, pp. 1457–1464. DOI: 10.1109/ICCV.2013.184.

[126] J. Yang, H. Li, D. Campbell, and Y. Jia. "Go-ICP: A Globally Optimal Solution to 3D ICP Point-Set Registration". In: *IEEE Transactions on Pattern*

Analysis and Machine Intelligence 38.11 (2016), pp. 2241–2254. DOI: 10.1109/TPAMI.2015.2513405.

[127] F. Bourier, A. Brost, L. Yatziv, J. Hornegger, N. Strobel, and K. Kurzidim. "Coronary Sinus Extraction for Multimodality Registration to guide Transseptal Puncture". In: *8th Interventional MRI Symposium*. Ed. by T. Kahn, F. A. Jolesz, and J. S. Lewin. Leipzig, 2010, pp. 311–313.

[128] A. Iskurt, Y. Becerikli, and K. Mahmutyazicioglu. "A Fast and Automatic Calibration of the Projectory Images for 3D Reconstruction of the Branchy Structures". In: *47th Annual Conference on Information Sciences and Systems (CISS)*. 2013, pp. 1–6.

[129] M.-P. Jolly, C. Guetter, X. Lu, H. Xue, and J. Guehring. "Automatic Segmentation of the Myocardium in Cine MR Images Using Deformable Registration". In: *Statistical Atlases and Computational Models of the Heart - Imaging and Modelling Challenges*. Vol. 8896. SEPTEMBER. 2011, pp. 105–113. DOI: 10.1007/978-3-642-28326-0_26. arXiv: 9780201398298.

[130] M. Panayiotou, A. P. King, R. J. Housden, Y. Ma, M. Cooklin, M. O'Neill, J. Gill, C. A. Rinaldi, and K. S. Rhode. "A Statistical Method for Retrospective Cardiac and Respiratory Motion Gating of Interventional Cardiac X-Ray Images". In: *Medical Physics* 41.7 (2014), p. 071901. DOI: 10.1118/1.4881140.

[131] E. Nasr-Esfahani, S. Samavi, N. Karimi, S. Soroushmehr, K. Ward, M. Jafari, B. Felfeliyan, B. Nallamothu, and K. Najarian. "Vessel extraction in X-ray angiograms using deep learning". In: *38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2016, pp. 643–646. DOI: 10.1109/EMBC.2016.7590784.

[132] S.-Y. Sun, P. Wang, S. Sun, and T. Chen. "Model-Guided Extraction of Coronary Vessel Structures in 2D X-Ray Angiograms". In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2014*. Ed. by P. Golland, N. Hata, C. Barillot, J. Hornegger, and R. Howe. Vol. 8674. Boston: Springer, 2014, pp. 594–602. DOI: 10.1007/978-3-319-10470-6_74.

[133] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. a. Viergever. "Multiscale vessel enhancement filtering". In: *Medial Image Computing and Computer-Assisted Invervention - MICCAI'98. Lecture Notes in Computer Science, vol 1496*. 1998, pp. 130–137. DOI: 10.1016/j.media.2004.08.001.

[134] Y. S. Chen and W. H. Hsu. "A Modified Fast Parallel Algorithm for Thinning Digital Patterns". In: *Pattern Recognition Letters* 7.2 (1988), pp. 99–106. DOI: 10.1016/0167-8655(88)90124-9.

[135] U. Ramer. "An Iterative Procedure for the Polygonal Approximation of Plane Curves". In: *Computer Graphics and Image Processing* 1.3 (1972), pp. 244–256. DOI: 10.1016/S0146-664X(72)80017-0.

[136] D. H. Douglas and T. K. Peucker. "Algorithms for the Reduction of the Number of Points Required to Represent a Digitized Line or its Caricature". In: *Cartographica: The International Journal for Geographic Information and Geovisualization* 10.2 (1973), pp. 112–122. DOI: 10.3138/FM57-6770-U75U-7727.

[137] R. J. Dakin. "A Tree-Search Algorithm for Mixed Integer Programming Problems". In: *The Computer Journal* 8.3 (1965), pp. 250–255. DOI: 10.1093/comjnl/8.3.250.

[138] A. H. Land and A. G. Doig. "An Automatic Method for Solving Discrete Programming Problems". In: *Econometrica* 28.3 (1960), pp. 497–520. DOI: 10.2307/1910129.

[139] D. Chetverikov and D. Stepanov. "Robust Euclidean Alignment of 3D Point Sets". In: *Image and Vision Computing* 23.3 (2005), pp. 299–309. DOI: 10.1016/j.imavis.2004.05.007.

[140] N. R. Van de Veire, J. D. Schuijf, J. De Sutter, D. Devos, G. B. Bleeker, A. de Roos, E. E. van der Wall, M. J. Schalij, and J. J. Bax. "Multimodality Imaging of Anatomy and Function in Coronary Artery Disease". In: *Journal of American College of Cardiology* 48.9 (2006), pp. 1832–1838. DOI: 10.1016/j.jacc.2006.07.042.

[141] M. Kaiser, M. John, A. Borsdorf, P. Mountney, R. Ionasec, A. Nöttling, P. Kiefer, and T. Neumuth. "Significant Acceleration of 2D-3D Registration-based Fusion of Ultrasound and X-ray Images by Mesh-based DRR Rendering". In: *SPIE Medical Imaging*. 2013, pp. 867111–867111.

[142] J. Kruger and R. Westermann. "Acceleration Techniques for GPU-based Volume Rendering". In: *Proceedings of the 14th IEEE Visualization 2003 (VIS'03)*. 2003, p. 38.

[143] S. G. Armato III, G. McLennan, L. Bidaut, M. F. McNitt-Gray, C. R. Meyer, A. P. Reeves, and L. P. Clarke. *Data From LIDC-IDRI*. 2015. DOI: 10.7937/K9/TCIA.2015.LO9QL9SX.

[144] G. Hinton, N. Srivastava, and K. Swersky. *Neural networks for machine learning lecture 6a overview of mini-batch gradient descent*. 2012.

[145] K. He, X. Zhang, S. Ren, and J. Sun. "Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification". In: *2015 IEEE International Conference on Computer Vision (ICCV)*. Santiago, Chile: IEEE, Dec. 2015, pp. 1026–1034. DOI: 10.1109/ICCV.2015.123.

[146] T. De Silva, A. Uneri, M. D. Ketcha, S. Reaungamornrat, G. Kleinszig, S. Vogt, N. Aygun, S.-F. Lo, J.-P. Wolinsky, and J. H. Siewerdsen. "3D–2D image registration for target localization in spine surgery: investigation of similarity metrics providing robustness to content mismatch". In: *Physics in medicine and biology* 61.8 (2016), pp. 3009–3025. DOI: 10.1088/0031-9155/61/8/3009..

[147] D. B. Russakoff, T. Rohlfing, K. Mori, D. Rueckert, A. Ho, J. R. Adler, and C. R. Maurer. "Fast Generation of Digitally Reconstructed Radiographs Using Attenuation Fields With Application to 2D-3D Image Registration". In: *IEEE Transactions on Medical Imaging* 24.11 (2005), pp. 1441–1454. DOI: 10.1109/TMI.2005.856749.

[148] T. Heimann, P. Mountney, M. John, and R. Ionasec. "Real-time ultrasound transducer localization in fluoroscopy images by transfer learning from synthetic training data". In: *Medical Image Analysis* 18.8 (2014), pp. 1320–1328. DOI: 10.1016/j.media.2014.04.007.

[149] Z. Tu. "Probabilistic Boosting-Tree: Learning Discriminative Models for Classification, Recognition, and Clustering". In: *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*. Beijing, China: IEEE, 2005, pp. 1589–1596. DOI: 10.1109/ICCV.2005.194.

[150] Y. Zhang, S. Miao, T. Mansi, and R. Liao. "Task driven generative modeling for unsupervised domain adaptation: Application to X-ray image segmentation". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 11071 LNCS. 2018, pp. 599–607. DOI: 10.1007/978-3-030-00934-2_67. arXiv: 1806.07201.

[151] J. Y. Zhu, T. Park, P. Isola, and A. A. Efros. "Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks". In: *Proceedings of the IEEE International Conference on Computer Vision*. Vol. 2017-Octob. 2017, pp. 2242–2251. DOI: 10.1109/ICCV.2017.244. arXiv: 1703.10593.

[152] M. Unberath, J. N. Zaech, S. C. Lee, B. Bier, J. Fotouhi, M. Armand, and N. Navab. "DeepDRR – A Catalyst for Machine Learning in Fluoroscopy-Guided Procedures". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 11073 LNCS. 2018, pp. 98–106. DOI: 10.1007/978-3-030-00937-3_12. arXiv: 1803.08606.

[153] J. Tobin, R. Fong, A. Ray, J. Schneider, W. Zaremba, and P. Abbeel. "Domain Randomization for Transferring Deep Neural Networks from Simulation to the Real World". In: *IEEE International Conference on Intelligent Robots and Systems*. 2017, pp. 23–30. DOI: 10.15607/RSS.2017.XIII.034. arXiv: 1611.04201.

[154] J. Tremblay, A. Prakash, D. Acuna, M. Brophy, V. Jampani, C. Anil, T. To, E. Cameracci, S. Boochoon, and S. Birchfield. "Training Deep Networks with Synthetic Data: Bridging the Reality Gap by Domain Randomization". In: *arXiv*. 2018, pp. 1–9. DOI: 10.1109/CVPRW.2018.00143. arXiv: 1804.06516.

[155] X. B. Peng, M. Andrychowicz, W. Zaremba, and P. Abbeel. "Sim-to-Real Transfer of Robotic Control with Dynamics Randomization". In: *arXiv* (2018), p. 1. DOI: 10.1007/s00267-013-0043-7. arXiv: arXiv:1710.06537v3.

[156] R. E. Kalman. "A New Approach to Linear Filtering and Prediction Problems1". In: *Journal of Basic Engineering* 82.1 (1960), pp. 35–45. DOI: 10.1115/1.3662552.

[157] P. Del Moral. "Nonlinear filtering: Interacting particle resolution". In: *Comptes Rendus de l'Académie des Sciences - Series I - Mathematics* 325.6 (1997), pp. 653–658. DOI: 10.1016/S0764-4442(97)84778-7.